

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 03:29:26 ; Search time 69.5142 Seconds
(without alignments)
3262.936 Million cell updates/sec

Title: US-09-996-617-2

Perfect score: 7534

Sequence: 1 MAGGAWGRLACYLEFLKKEE.....HLIMELWEKSKGLPLSS 1429

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: A_Geneseq_19Jun03.*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7534	100.0	1429	23	Human CARD-7 polyp
2	7534	100.0	1429	22	Human leucine rich
3	7534	100.0	1429	23	Leucine-rich repea
4	7534	100.0	1429	23	Human leucine-rich
5	7534	100.0	1429	24	Human caspase recr
6	7534	100.0	1429	24	Human caspase recr
7	7502	99.6	1473	22	Human NB-ARC and C
8	7502	99.6	1473	22	Human NAC beta iso
9	7488	99.4	1429	23	Human caspase recr

10	7488	99.4	1429	23	AAO17855	Pyrin domain conta
11	7484	99.3	1473	22	AAE06758	Human G-protein co
12	7342	97.5	1397	22	AAO72670	Human NB-ARC and C
13	7325.5	97.2	1442	22	AAO72671	Human NB-ARC and C
14	6344	84.2	1454	22	AAO72673	Human NAC beta iso
15	6174	81.9	1424	22	AAO72674	Human NAC gamma or
16	2171.5	28.8	442	21	AAO24513	Human secreted pro
17	1380	18.3	1061	23	AAO15590	Human PYRIN-8 prot
18	1375	18.3	1099	23	AAO17857	Pyrin domain conta
19	1279	17.0	1034	22	AAO7514	Human PYRIN-1 prot
20	1279	17.0	1034	24	ABU08503	Huma PYRIN-1 prote
21	1263.5	16.8	920	23	ABP53254	Human MDT-13 prot
22	1154	15.3	896	23	AAO15592	Human PYRIN-11 prot
23	1037	13.8	994	23	AAO15593	Human PYRIN-3 prot
24	1020.5	13.5	983	23	AAO17870	Pyrin domain conta
25	1007	13.4	919	23	ABP43483	Human secreted pro
26	1007	13.4	919	24	ABJ19319	NOVX related prote
27	985.5	13.1	1162	23	AAE25053	Human EMERY-1 prot
28	983	13.0	1200	23	AAU79526	Human MATER protei
29	983	13.0	1200	24	AAE31749	Human MATER protei
30	982.5	13.0	2312	23	AAO17859	Pyrin domain conta
31	972.5	12.9	858	23	AAO50328	Human nucleotide b
32	948	12.6	1062	22	AAO65895	Amino acid sequenc
33	948	12.6	1062	23	AAO17856	Pyrin domain conta
34	945.5	12.5	1033	22	AAO7513	Human nucleotide b
35	945.5	12.5	1033	23	ABG78456	Human nucleotide b
36	945.5	12.5	1033	24	ABU08502	Huma pyrin domain
37	945.5	12.5	1344	23	AAO15585	Human PYRIN-5 prot
38	941.5	12.5	1049	22	AAO65894	Amino acid sequenc
39	934.5	12.4	1033	23	ABU65214	Human NOV125b prot
40	924.5	12.3	732	23	AAO17858	Pyrin domain conta
41	922.5	12.2	952	22	ABG03924	Novel human diagno
42	910	12.1	1022	23	AAO17862	Pyrin domain conta
43	908.5	12.1	980	23	ABB98244	Human CGPD SQ ID
44	908.5	12.1	980	23	AAO50329	Human nucleotide b
45	897	11.9	891	22	AAO45446	Human G-protein co

ALIGNMENTS

RESULT 1
AAB62571
ID AAB62571 standard; Protein; 1429 AA.

XX	AC	AAO62571;			
XX	XX	23-JUL-2001 (first entry)			
DT	XX	Human CARD-7 polypeptide.			
DE	XX	CARD-7; CARD-8; CARD-5; caspase recruitment domain; cancer; human;			
KW	KW	autoimmune disorder; antiinflammatory; immunosuppressive; antiallergic;			
KW	KW	antibacterial; antiviral; gene therapy.			
XX	OS	Homo sapiens.			
XX	XX	Key	Location/Qualifiers		
FT	FT	Domain	329..645	/note="nucleotide binding domain"	
FT	FT	Domain	333..341	/note="kinase 1A subdomain"	
FT	FT	Domain	404..413	/note="kinase 2 subdomain"	
FT	FT	Domain	454..463	/note="kinase 3a subdomain"	
FT	FT	Domain	615..622	/note="motif 2 domain"	
FT	FT	Domain	807..834	/note="leucine-rich domain"	
FT	FT	Domain	836..863	/note="leucine-rich domain"	
FT	FT	Domain	864..891		

FT Domain /note= "leucine-rich domain"
PT 893..920
FT /note= "leucine-rich domain"
PT 921..948
FT /note= "leucine-rich domain"
PT 950..976
FT /note= "leucine-rich domain"
XX WO200130813-A1.
XX 03-MAY-2001.
XX 27-OCT-2000; 2000WO-US29796.
XX 27-OCT-1999; 99US-0428252.
XX (MILL-) MILLENNIUM PHARM INC.
XX Bertin J;
XX WPI; 2001-343340/36.
XX N-PSDB; AAF83651.
XX
PT Identifying a modulator of interaction between caspase recruitment
PT domain (CARD)-7 and CARD-5, for treating autoimmune disorders,
PT comprises measuring the binding of CARD-7 and CARD-5 in the presence of
PT the compound -
XX
PS Disclosure; Fig 1A-D; 80pp; English.
XX
CC The invention relates to identifying compounds, that modulate interaction
CC between caspase recruitment domain (CARD)-7 and CARD-5. The method
CC involves measuring the binding of CARD-7 and CARD-5 in the presence of
CC the compound (an increase in the binding of CARD-7 to CARD-5 in the
CC presence of the compound compared to the binding in the absence of the
CC compound indicates that the compound is a modulator of CARD-7-CARD-5
CC interaction). Modulators of CARD-7 and CARD-8 expression or activity can
CC be used to treat or diagnose disorders such as cancers, bacterial or
CC viral infections, autoimmune disorders (systemic lupus erythematosus,
CC immune-mediated glomerulonephritis or arthritis), inflammatory disorders,
CC organ-specific autoimmunity including multiple sclerosis, Hashimoto's
CC thyroiditis, or Grave's disease, psoriasis, graft rejection, allergies.
CC CARD-7 and CARD-8 are useful as modulating agents in regulating a variety
CC of cellular processes including cell growth and cell death. The present
CC sequence represents the human CARD-7, an intracellular protein.
XX
SQ Sequence 1429 AA;
Query Match 100.0%; Score 7534; DB 22; Length 1429;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAGGAWGLACYLEFLKKEELKEFOLLANKAHSRSSGGETPAQPEKTSMEVASYLVAQ 60
DB 1 MAGGAWGLACYLEFLKKEELKEFOLLANKAHSRSSGGETPAQPEKTSMEVASYLVAQ 60
QY 61 YGEQAWDLALHTWQMGRLSLCAQOAGAGHSPSPFSPHGLSPSQPTSTAVLMPW 120
DB 61 YGEQAWDLALHTWQMGRLSLCAQOAGAGHSPSPFSPHGLSPSQPTSTAVLMPW 120
QY 121 IHELPAQCTQGSERRVLRQLPTSGRRWREISASLLYQALPSSPDHESPSQSPNAPTST 180
DB 121 IHELPAQCTQGSERRVLRQLPTSGRRWREISASLLYQALPSSPDHESPSQSPNAPTST 180
QY 181 AVLGSWGPQPSLAPREQEAFGTQWPLDETSGIYYTEIREREREKSEKGRPPAAVVG 240
DB 181 AVLGSWGPQPSLAPREQEAFGTQWPLDETSGIYYTEIREREREKSEKGRPPAAVVG 240
QY 241 PPQAHTSLQPHHPHWPSPVRESLCSTWPKWNEFQKFTQLLLLRPHRPSQDPLVKRSW 300
DB 241 PPQAHTSLQPHHPHWPSPVRESLCSTWPKWNEFQKFTQLLLLRPHRPSQDPLVKRSW 300
QY 301 PDYVEENRGHLIEIRDLFGPGLDTQEPRIVLQGAAGIGKSTLARQVKEAWGRGQLYGR 360

301 PDYVEENRGHLIEIRDLFGPGLDTQEPRIVLQGAAGIGKSTLARQVKEAWGRGQLYGR 360
361 FQHVFPSCRELAQSKVSLAELIGKOGTATPAPIRQILSRPERLLFLDGDVDEPGWLQ 420
361 FQHVFPSCRELAQSKVSLAELIGKOGTATPAPIRQILSRPERLLFLDGDVDEPGWLQ 420
421 EPSSELCLHWSQOPADALGSLGKTLIPASFLITARTTALQNLIPSLQARWVEVLG 480
421 EPSSELCLHWSQOPADALGSLGKTLIPASFLITARTTALQNLIPSLQARWVEVLG 480
481 FSESSEKRYFYRYFTDERQAIAPRLVKSXKELMALCLVPWVSMALACTCLQOMKREKL 540
481 FSESSEKRYFYRYFTDERQAIAPRLVKSXKELMALCLVPWVSMALACTCLQOMKREKL 540
541 TLTSKTTTTLCHLYLAQALQOPIGPQRLDCLSLAAGIWQKTLFSPDDLRKHGLDGA 600
541 TLTSKTTTTLCHLYLAQALQOPIGPQRLDCLSLAAGIWQKTLFSPDDLRKHGLDGA 600
601 ISTFLKMGILQEHPIPLSYFIHLCFQEFFAAMSYLEDEKGRGHSHNCIIDLKTL 660
601 ISTFLKMGILQEHPIPLSYFIHLCFQEFFAAMSYLEDEKGRGHSHNCIIDLKTL 660
661 GIHGLFGASTTRFLGLLSDSEGEREMENIPHCRLSQGRNLMQWVPSLQLLQPHSLSLH 720
661 GIHGLFGASTTRFLGLLSDSEGEREMENIPHCRLSQGRNLMQWVPSLQLLQPHSLSLH 720
721 CLYETRNKTFLTQWMAHFEENGMCVETDMLLVCTFCIKFSRHVKQLQIIEGRQHRSTWS 780
721 CLYETRNKTFLTQWMAHFEENGMCVETDMLLVCTFCIKFSRHVKQLQIIEGRQHRSTWS 780
781 PTMVLFRWVPTDAYWOLIFSVLKVTRNLKELDLSGNSLSHSAVSKLCTLRPRCLLE 840
781 PTMVLFRWVPTDAYWOLIFSVLKVTRNLKELDLSGNSLSHSAVSKLCTLRPRCLLE 840
841 TRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLQRLRQPSCKLRLQ 900
841 TRLAGCGLTAEDCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLQRLRQPSCKLRLQ 900
901 LVSCGLTSDCCQDLASVLSASPSLKDLOQNNLDVGVRLLCGLRHPACKLRLGLDQ 960
901 LVSCGLTSDCCQDLASVLSASPSLKDLOQNNLDVGVRLLCGLRHPACKLRLGLDQ 960
961 TTLSDEMROELRALOEKPOLLIIFSRKXPSVMTTEGLDTGEMNSSTSLKRLQSGSRA 1020
961 TTLSDEMROELRALOEKPOLLIIFSRKXPSVMTTEGLDTGEMNSSTSLKRLQSGSRA 1020
1021 ASHVAQANLKLLDVSKIPPIAIEAESSESPVVPVELLCVPSPASQGLHTKPLGTDDDFW 1080
1021 ASHVAQANLKLLDVSKIPPIAIEAESSESPVVPVELLCVPSPASQGLHTKPLGTDDDFW 1080
1081 GPTGPVATEVVDKKNLYRVHPFVAGSYRWNTGLCFVMREAVTVEIFPCWQDQPLGIN 1140
1081 GPTGPVATEVVDKKNLYRVHPFVAGSYRWNTGLCFVMREAVTVEIFPCWQDQPLGIN 1140
1141 PQHSMWVAGPLDITKABPGAVEAHLPHFVALQGHVDTSLFQMAHFKEGMLEKPARV 1200
1141 PQHSMWVAGPLDITKABPGAVEAHLPHFVALQGHVDTSLFQMAHFKEGMLEKPARV 1200
1201 ELHIVLENPSFPLGVLKMIHNAIRPIVTSVVLVYHVRHPEVTPHLYLIPSDCSIR 1260
1201 ELHIVLENPSFPLGVLKMIHNAIRPIVTSVVLVYHVRHPEVTPHLYLIPSDCSIR 1260
1261 KELELCYRSPGEDQLFSEFYVGHGSGIRLQVKDKOETLWELVWALVXPGDMLPATLIPP 1320
1261 KELELCYRSPGEDQLFSEFYVGHGSGIRLQVKDKOETLWELVWALVXPGDMLPATLIPP 1320
1321 ARIAVPSPLDAPQLHFDVQVREOLIAVTSVEVVLKHLGQVLSQOEYERVLANTRPS 1380
1321 ARIAVPSPLDAPQLHFDVQVREOLIAVTSVEVVLKHLGQVLSQOEYERVLANTRPS 1380
1381 QMRKLFSLSQSWDRCKCKDGLYQALKETHPHILMELWEKSGSKGLPLSS 1429

Db 1381 QMRKLFSLQSQWDRKCKDGLYOALKETHPLHLMELWEKSGKGLPLSS 1429

RESULT 2
ABG97969
ID ABG97969 standard; Protein; 1429 AA.
XX
AC ABG97969;
XX
DT 07-JAN-2003 (first entry)
XX
DE Human leucine rich repeat domain containing protein #2.
XX
KW Leucine rich repeat; nervous system; human; neural disorder; apoptosis;
KW renal disorder; immune disorder; arthritis; asthma; AIDS;
KW acquired immunodeficiency syndrome; rheumatoid arthritis;
KW haematopoietic disorder; metabolic disorder; reproductive disorder;
KW pulmonary disease; cardiovascular disease; hyperproliferative disorder;
KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
KW Huntington's disease; developmental disorder; autoimmune disease;
KW Addison's disease; haemolytic anaemia; antiphospholipid syndrome;
KW allergic encephalomyelitis; gene therapy.

XX Homo sapiens.

XX WO200274959-A2.

XX 26-SEP-2002.

PF 20-DEC-2001; 2001WO-US50457.

PR 03-JAN-2001; 2001US-259479P.

PR 09-JAN-2001; 2001US-260616P.

XX (BRIM) BRISTOL-MYERS SQUIBB CO.

XX Ramanathan C, Feder J, Mintier G;

XX WPI; 2002-750554/81.

XX New HLRNS1 nucleic acids and polypeptides, useful for preventing,
PT treating, or ameliorating e.g. renal disorder, immune, hematopoietic,
PT metabolic, reproductive, pulmonary, cardiovascular or autoimmune
PT diseases

XX Example 1; Page 392-397; 415pp; English.

XX The invention describes nucleic acids encoding human leucine-rich repeat
CC containing proteins expressed in nervous system tissues, HLRNS1. The
CC HLRNS1 polypeptide or the polynucleotide is useful for preventing,
CC treating, or ameliorating a neural disorder or a disorder related to
CC aberrant apoptosis modulation (either directly or indirectly), renal
CC disorder, immune disorder (e.g. arthritis, asthma, acquired
CC immunodeficiency syndrome (AIDS) or rheumatoid arthritis),
CC haematopoietic, metabolic, reproductive, pulmonary or cardiovascular
CC diseases, hyperproliferative disorders, neurodegenerative diseases
CC (e.g. Alzheimer's disease, Parkinson's disease or Huntington's
CC disease), developmental disorders, non-infections disorders, nervous
CC system diseases and/or disorders, and autoimmune diseases (e.g.
CC Addison's disease, haemolytic anaemia, antiphospholipid syndrome, or
CC allergic encephalomyelitis). The polynucleotides are also useful as
CC chromosome markers, for chromosome identification, gene therapy, and in
CC identifying organisms from minute biological samples. This is the amino
CC acid sequence of a leucine-rich repeat containing protein.

XX Sequence 1429 AA;

Query Match 100.0%; Score 7534; DB 23; Length 1429;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGGAWGRLCYLEFLKKEBELKFOLLANKAHSRSSGGETPAQPEKTSGMVEASYLVAQ 60
|||||

Db 1 MAGGAWGRLCYLEFLKKEBELKFOLLANKAHSRSSGGETPAQPEKTSGMVEASYLVAQ 60
QY 61 YGEORAWDLALHTWEQMGRLSLCAQAQAGAGHSFPYSPSEPHLGSPSQPTSTAVLMPW 120
Db 61 YGEORAWDLALHTWEQMGRLSLCAQAQAGAGHSFPYSPSEPHLGSPSQPTSTAVLMPW 120
QY 121 IHELPAAGCTQGSERRVLRQLPDTSGRRWRREISASLLYQALPSSPDHSPSQESPNAPTST 180
Db 121 IHELPAAGCTQGSERRVLRQLPDTSGRRWRREISASLLYQALPSSPDHSPSQESPNAPTST 180
QY 181 AVLGSWGSPPPSLAPREQAPGTQWPLDETSGIYYTEIREREREKSKGKPPWAAVGT 240
Db 181 AVLGSWGSPPPSLAPREQAPGTQWPLDETSGIYYTEIREREREKSKGKPPWAAVGT 240
QY 241 PPQAHTSLQPHHHHPWPSVRESLCSCTWPKNEDFNQKFTQLLLQRPSPRQDPLVKESW 300
Db 241 PPQAHTSLQPHHHHPWPSVRESLCSCTWPKNEDFNQKFTQLLLQRPSPRQDPLVKESW 300
QY 301 PDYVEENRGHLIEIRDLFGPLDQEPRIIVLOGAAGIGKSTLARQVKEAWRGQLYGDR 360
Db 301 PDYVEENRGHLIEIRDLFGPLDQEPRIIVLOGAAGIGKSTLARQVKEAWRGQLYGDR 360
QY 361 FQHVYFSCRELAQSKVSLAELIGKGTATPAPIRQILSRPERLLFLDGVDEPGWLQ 420
Db 361 FQHVYFSCRELAQSKVSLAELIGKGTATPAPIRQILSRPERLLFLDGVDEPGWLQ 420
QY 421 EPSSELCLHWSQOPADALLGSLGKTLIPASFLITARTTALQNLIPSEQARVVEVLG 480
Db 421 EPSSELCLHWSQOPADALLGSLGKTLIPASFLITARTTALQNLIPSEQARVVEVLG 480
QY 481 FSESRSKEYFYRYFTDERQAIRAFRLVKSNEKELWALCLVPVWVSLAECTCLMQMKRKEKL 540
Db 481 FSESRSKEYFYRYFTDERQAIRAFRLVKSNEKELWALCLVPVWVSLAECTCLMQMKRKEKL 540
QY 541 TLTSKTTTTLCLHYLAQALQAPLGPQPLDCLSLAAGIWKQKTLFSDDLRKHGDGAI 600
Db 541 TLTSKTTTTLCLHYLAQALQAPLGPQPLDCLSLAAGIWKQKTLFSDDLRKHGDGAI 600
QY 601 1STFLKMGILQEHPIPLSYSFHLCFOFFAAMSYYLDEDEKRGKHSNCIIDEKTLLEY 660
Db 601 1STFLKMGILQEHPIPLSYSFHLCFOFFAAMSYYLDEDEKRGKHSNCIIDEKTLLEY 660
QY 661 GIHGLFGASTTRFLGLLSDGEGEMENIFHCRLSQGRNLMQWVPSLQLLQPHSLESIH 720
Db 661 GIHGLFGASTTRFLGLLSDGEGEMENIFHCRLSQGRNLMQWVPSLQLLQPHSLESIH 720
QY 721 CLYETRNKTLTQVMAHFEEMGCMVETDMLLVCTFCIKFSRHVKKLOLIEGRQHRSTWS 780
Db 721 CLYETRNKTLTQVMAHFEEMGCMVETDMLLVCTFCIKFSRHVKKLOLIEGRQHRSTWS 780
QY 781 PTMVVLFWRVPVTDAYWQILFSLVLRNLKELDLSGNSLSHSAVKSCKTLRRPRCLLE 840
Db 781 PTMVVLFWRVPVTDAYWQILFSLVLRNLKELDLSGNSLSHSAVKSCKTLRRPRCLLE 840
QY 841 TLRLACGLTAPDCXDLAFGLRANQTLTDLDSFNVLTDAGAKHLCQRLQPSCKLQRIQ 900
Db 841 TLRLACGLTAPDCXDLAFGLRANQTLTDLDSFNVLTDAGAKHLCQRLQPSCKLQRIQ 900
QY 901 LVSCGLTSDCCODLASVLSASPSLKELDLQNNLDDVGRLLCEGHRHPACKLIRGLDQ 960
Db 901 LVSCGLTSDCCODLASVLSASPSLKELDLQNNLDDVGRLLCEGHRHPACKLIRGLDQ 960
QY 961 TTLSDEMQEQLRALQEQKPOLLI FSRRKPSVMTPTTEGLDTGEMSNSTSLKQRLGSERA 1020
Db 961 TTLSDEMQEQLRALQEQKPOLLI FSRRKPSVMTPTTEGLDTGEMSNSTSLKQRLGSERA 1020
QY 1021 ASHVAQANLKLLDVSKIFPIAEIABESSPEVVPVELLCVPSPASQGLHTKPLGTDGDFW 1080
Db 1021 ASHVAQANLKLLDVSKIFPIAEIABESSPEVVPVELLCVPSPASQGLHTKPLGTDGDFW 1080
QY 1081 GPTGPVATEVDVKENKLYRVHPFPVAGSYRWPNVTGLCFVMREAVTVEIEFCVWDQFLGEIN 1140
Db 1081 GPTGPVATEVDVKENKLYRVHPFPVAGSYRWPNVTGLCFVMREAVTVEIEFCVWDQFLGEIN 1140

QY 1141 POHSMVAGPLLDIKAEPCGAVAVHLPHPFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
 DB 1141 POHSMVAGPLLDIKAEPCGAVAVHLPHPFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
 QY 1201 ELHHIVLENPSPPLGVLLKMIHNAURFIPVTSVLLYHRVHPPEVTPHLYLIPSDCSIR 1260
 DB 1201 ELHHIVLENPSPPLGVLLKMIHNAURFIPVTSVLLYHRVHPPEVTPHLYLIPSDCSIR 1260
 QY 1261 KELELCYRSPGEDOLFSEBYVGHLSGIRLQVKKDETLVWEALVKPDGLMPATTLIPP 1320
 DB 1261 KELELCYRSPGEDOLFSEBYVGHLSGIRLQVKKDETLVWEALVKPDGLMPATTLIPP 1320
 QY 1321 ARIAVPSLPDAPOLLHFVQYRQLIARTVTSVEVLDLKHGOVLQSOEQYERVLAEINTPS 1380
 DB 1321 ARIAVPSLPDAPOLLHFVQYRQLIARTVTSVEVLDLKHGOVLQSOEQYERVLAEINTPS 1380
 QY 1381 QMRKLFSLQSQSDRKCKDGLYQALKETHPHLIMELWEKSKGLPLSS 1429
 DB 1381 QMRKLFSLQSQSDRKCKDGLYQALKETHPHLIMELWEKSKGLPLSS 1429

RESULT 3
 ID ABG78472
 XX ABG78472 standard; Protein; 1429 AA.

AC ABG78472;
 XX

DT 15-NOV-2002 (first entry)

DE Leucine-rich repeat (LRR) domain containing protein #2.

KW Human; human leucine-rich repeat small intestine I; HLRRS11; asthma;
 KW proliferative disorder; gastrointestinal disorder; renal disorder;
 KW neural disorder; reproductive disorder; calcium regulation; apoptosis;
 KW immune system; anemia; human immune deficiency virus; HIV; cancer;
 KW blood coagulation disorder; autoimmune disorder; allergic reaction;
 KW inflammatory condition; cardiovascular disorder; ischaemia;
 KW neurological disorder; infectious disease; cytokine production;
 KW expressed sequence tag; EST.

OS Unidentified.

OS WO200261086-A2.

PN 08-AUG-2002.

XX 20-DEC-2001; 2001WO-US49739.

XX 22-DEC-2000; 2000US-257774P.

XX (BRIM) BRISTOL-MYERS SQUIBB CO.

XX Feder J, Ramanathan C, Mintier G;

XX WPI; 2002-619252/66.

XX New isolated nucleic acid molecules encoding HLRRS11 polypeptides, or
 PT their fragments and homologues, useful for preventing, treating and
 PT ameliorating medical conditions, e.g. proliferative, gastrointestinal,
 PT or renal disorders -

XX Disclosure, Page 229-233; 336pp; English.

XX The invention relates to isolated nucleic acid molecules (I) encoding
 CC human leucine-rich repeat small intestine I (HLRRS11) polypeptides.
 CC The nucleic acid molecules and polypeptides are useful for preventing,
 CC treating and ameliorating medical conditions, such as proliferative,
 CC gastrointestinal, renal, neural, or reproductive disorders; or disorders
 CC related to aberrant calcium regulation or apoptosis modulation, either
 CC directly or indirectly. They are also useful for treating, preventing
 CC and/or diagnosing diseases, disorders and/or conditions of: immune system
 CC by activating or inhibiting the proliferation, differentiation, or

CC mobilisation of immune cells; haematopoietic cells e.g. thrombocytopenia,
 CC anaemia; immunologic deficiency syndromes, e.g. human immune deficiency
 CC virus (HIV) infection, HTLV-BIV infection; blood coagulation disorders,
 CC e.g. arterial thrombosis; autoimmune disorders, e.g. Addison's disease,
 CC myasthenia gravis; asthma or allergic reactions; inflammatory
 CC conditions, e.g. chronic prostatitis, sepsis; proliferative disorders,
 CC e.g. cancer; cardiovascular disorders, e.g. arrhythmia, myocardial
 CC ischaemias, aneurysms; neurological disorders, e.g. Alzheimer's disease,
 CC Huntington's chorea; infectious diseases, e.g. measles, mumps,
 CC pneumonia, or viral, bacterial, and fungal infections. The HLRRS11
 CC polypeptides are useful for modulating cytokine production, antigen
 CC presentation, or other processes such as boosting immune responses.
 CC ABG78454-ABG78474 represent HLRRS11 amino acid sequences and related
 CC amino acid sequences of the invention.

XX Sequence 1429 AA;

Query Match 100.0%; Score 7534; DB 23; Length 1429;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGGAWGLACYLEFLKKEELKEFOLLANKAHSSSGGETPAQPEKTSQGEVASYLVAQ 60
 DB 1 MAGGAWGLACYLEFLKKEELKEFOLLANKAHSSSGGETPAQPEKTSQGEVASYLVAQ 60
 QY 61 YGEQAWDLALHTWEQMGRLSLCAQAGAGHSPSPYSPSEPHLGSPSQPTSTAVLMPW 120
 DB 61 YGEQAWDLALHTWEQMGRLSLCAQAGAGHSPSPYSPSEPHLGSPSQPTSTAVLMPW 120
 QY 121 IHELPAGCTQGSERRVLRQLPDTSGRRWRREISASLLYQALPSSPDHSPSPESNAPTST 180
 DB 121 IHELPAGCTQGSERRVLRQLPDTSGRRWRREISASLLYQALPSSPDHSPSPESNAPTST 180
 QY 181 AVLGSWGSPPOPSLAPRQEPAGTQWPLDETSGIYYTIREREREKSKGPPHAAVGT 240
 DB 181 AVLGSWGSPPOPSLAPRQEPAGTQWPLDETSGIYYTIREREREKSKGPPHAAVGT 240
 QY 241 PQAHTSLQPHHPWPESVRESLCSLTPWKNEFNQKFTQLLLQRPHPRQDPLVKRSW 300
 DB 241 PQAHTSLQPHHPWPESVRESLCSLTPWKNEFNQKFTQLLLQRPHPRQDPLVKRSW 300
 QY 301 PDYVEENRGHLIEIRDLPGLDTQEPRIVILQGAAGIGKSTLARQVKEAMRGOLYGD 360
 DB 301 PDYVEENRGHLIEIRDLPGLDTQEPRIVILQGAAGIGKSTLARQVKEAMRGOLYGD 360
 QY 361 FQHVYFSCRELAQSKVVSLELIGKGTATPAPIRQILSPERLLFLDGVDEPGWVLQ 420
 DB 361 FQHVYFSCRELAQSKVVSLELIGKGTATPAPIRQILSPERLLFLDGVDEPGWVLQ 420
 QY 421 EPSSELCLHWSQOPADALLGLGKTLTPASFLITARTTALQNLIPSLQARWVEVLG 480
 DB 421 EPSSELCLHWSQOPADALLGLGKTLTPASFLITARTTALQNLIPSLQARWVEVLG 480
 QY 481 FSESSEKEYFYRYFTDERQAIAPRLVKSNKELWALCLVPVWSMLACTCLQOMKREKL 540
 DB 481 FSESSEKEYFYRYFTDERQAIAPRLVKSNKELWALCLVPVWSMLACTCLQOMKREKL 540
 QY 541 TLTKTTTTLCLHYLAQALQAPLGPQLRDLCSLAAGIWKQKTLFSPDDLRKHGLDGI 600
 DB 541 TLTKTTTTLCLHYLAQALQAPLGPQLRDLCSLAAGIWKQKTLFSPDDLRKHGLDGI 600
 QY 601 ISTFLKMGILQEHPIPLSYFHLICFOFFFAAMSIVLEDEKRGHSHNCIIDLKTL 660
 DB 601 ISTFLKMGILQEHPIPLSYFHLICFOFFFAAMSIVLEDEKRGHSHNCIIDLKTL 660
 QY 661 GIHGLFGASTTRFLGLLSDGEREMENI FHCRLSQGRNLQMWVPSQLQLLQPHSLESILH 720
 DB 661 GIHGLFGASTTRFLGLLSDGEREMENI FHCRLSQGRNLQMWVPSQLQLLQPHSLESILH 720
 QY 721 CLYETRNKTFUTQWMAHFEEMGCMVETDMELLVCTFCIKFSRHVKKQLIIEGRQHRSTWS 780
 DB 721 CLYETRNKTFUTQWMAHFEEMGCMVETDMELLVCTFCIKFSRHVKKQLIIEGRQHRSTWS 780


```
QY 781 PTMVVLFRRVPTDAYWQILFVLKVTNRNLKELDLGNSLSHSAVKSCKTLRRPRCLLE 840
Db 781 PTMVVLFRRVPTDAYWQILFVLKVTNRNLKELDLGNSLSHSAVKSCKTLRRPRCLLE 840
QY 841 TLRLAGCGLTAECDCKDLAFGLRANOTLTFLDLS FNVLTDAKAKHLCOBLRQPSCKLQRLQ 900
Db 841 TLRLAGCGLTAECDCKDLAFGLRANOTLTFLDLS FNVLTDAKAKHLCOBLRQPSCKLQRLQ 900
QY 901 LVSCGLTSDCCQDLASVLSASPSLKELDLQNNLDDVGVRLCEGLRHPACKLIRLGLDQ 960
Db 901 LVSCGLTSDCCQDLASVLSASPSLKELDLQNNLDDVGVRLCEGLRHPACKLIRLGLDQ 960
QY 961 TTLSDEMROELRALBOEKPOLLI FRRKPSVMTPTTEGLDTGMSNSTSLKRRQLGSGRA 1020
Db 961 TTLSDEMROELRALBOEKPOLLI FRRKPSVMTPTTEGLDTGMSNSTSLKRRQLGSGRA 1020
QY 1021 ASHVAQANLKLDDVSKIFPIAIBAESSPEVVPVELLCVPSASQGDHLTKPLGTDDDFW 1080
Db 1021 ASHVAQANLKLDDVSKIFPIAIBAESSPEVVPVELLCVPSASQGDHLTKPLGTDDDFW 1080
QY 1081 GPTGPVATEVVDKEKNLYRVHPFPVAGSYRWPNTGLCFVMREAVTVIEIFCVWDQFLGEIN 1140
Db 1081 GPTGPVATEVVDKEKNLYRVHPFPVAGSYRWPNTGLCFVMREAVTVIEIFCVWDQFLGEIN 1140
QY 1141 PQHSMWVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
Db 1141 PQHSMWVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
QY 1201 ELHHIVLENPSPGLGVLLKMTNARLFPVTSVLLYHRVHPPEVTHLYLIPSDCSIR 1260
Db 1201 ELHHIVLENPSPGLGVLLKMTNARLFPVTSVLLYHRVHPPEVTHLYLIPSDCSIR 1260
QY 1261 KSELELCYRSPGDDQLFSEFVVGHLGGIRLQVKDKDELTVWEALVKPGDLMPATTLIPP 1320
Db 1261 KSELELCYRSPGDDQLFSEFVVGHLGGIRLQVKDKDELTVWEALVKPGDLMPATTLIPP 1320
QY 1321 ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLKLGQVLSQEQYERVLAENTRPS 1380
Db 1321 ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVLKLGQVLSQEQYERVLAENTRPS 1380
QY 1381 QMRKLFSLSQSWDRCKDGLYQALKETHPHLIMELWEKSKGGLPLSS 1429
Db 1381 QMRKLFSLSQSWDRCKDGLYQALKETHPHLIMELWEKSKGGLPLSS 1429

RESULT 4
ABB77916
ID ABB77916 standard; protein; 1429 AA.
XX ABB77916;
AC ABB77916;
XX ABB77916;
DT 07-OCT-2002 (first entry)
XX Human leucine-rich repeat domain containing protein KIAA0926.
DE Human; leucine-rich repeat; HLRRBM1; proliferative disorder;
XX immune condition; apoptosis; signal transduction; autoimmune disease;
KW haematopoietic cell disease; graft-versus-host disease; allergy; asthma;
KW cardiovascular disorder; neurological disease; pheromone;
KW pulmonary disease; bronchial obstructive pulmonary disease;
KW allergic rhinitis; chronic hypersensitivity; reproductive disease;
KW haematopoietic disease; platelet disorder; Bernard-Soulier
KW inflammatory disorder; systemic lupus erythematosus;
KW cardiovascular disease; cancer; KIAA0926.
XX Homo sapiens.
OS
XX WO200252011-A2.
PN
XX 04-JUL-2002.
PD
XX 20-DEC-2001; 2001WO-US49740.
PF
XX
```

```
PR 22-DEC-2000; 2000US-257773P.
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
PA Feder J, Ramanathan C, Mintier G;
XX WPI; 2002-566676/60.
XX New HLRRBM1 nucleic acids for preventing, treating or ameliorating e.g.
PT proliferative disorders, immune conditions, a disorder related to
PT aberrant apoptosis modulation or developmental disorders -
XX
XX Example 1; Page 362-366; 371pp; English.
XX
XX The present sequence represents a human leucine-rich repeat domain
CC containing protein, which was used as a probe to search for leucine-rich
CC repeat containing protein HLRRBM1. HLRRBM1 polypeptides and
CC polynucleotides are useful for preventing, treating or ameliorating a
CC medical condition such as a proliferative disorder, immune condition, or
CC a disorder related to aberrant apoptosis modulation, either directly or
CC indirectly, and in modulating signal transduction activity in various
CC cells, tissue and organisms. They are also useful for treating,
CC preventing, or diagnosing diseases of haematopoietic cells, autoimmune
CC disease, graft-versus-host disease, allergic conditions (e.g. asthma),
CC cardiovascular disorders, and neurological diseases, and for increasing
CC the organisms' ability to synthesize and/or release pheromones. The
CC polypeptide may also be used in treating, preventing or ameliorating
CC pulmonary disease (e.g. chronic obstructive pulmonary disease, allergic
CC rhinitis, or bronchial hyperresponsiveness), reproductive disease,
CC haematopoietic disease, platelet disorders (e.g. Bernard-Soulier
CC syndrome), non-infectious immune response), immune and inflammatory
CC disorders (e.g. systemic lupus erythematosus), cardiovascular diseases
CC and cancers. HLRRBM1 nucleic acids may further be used in chromosome
CC identification or mapping, as a chromosome marker, as molecular weight
CC markers, as diagnostic probes, in gene therapy, in raising anti-DNA
CC antibodies, or as antigens for eliciting immune responses.
XX
XX Sequence 1429 AA;
SQ
Query Match 100.0%; Score 7534; DB 23; Length 1429;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAGGAWGLACYLEFLKKEELKEFOLLANKAHSSSSGCTPAQPEKTSGMGEVASYLVAQ 60
Db 1 MAGGAWGLACYLEFLKKEELKEFOLLANKAHSSSSGCTPAQPEKTSGMGEVASYLVAQ 60
QY 61 YGEQRANDLALHTWEQMGRLSLCAQAQEGAGHSFPYSPSEPHLGSPSQPTSTAVLMPW 120
Db 61 YGEQRANDLALHTWEQMGRLSLCAQAQEGAGHSFPYSPSEPHLGSPSQPTSTAVLMPW 120
QY 121 IHELPAGTCQSERRVRLQPLDTSGRRRWRETSASLLYQALPSSPDHESPNAPTST 180
Db 121 IHELPAGTCQSERRVRLQPLDTSGRRRWRETSASLLYQALPSSPDHESPNAPTST 180
QY 181 AVLGSWGGPPPSLAPRQEQAPGTOWPLDTSIGIYITIRERERESKGRPPWAAVVG 240
Db 181 AVLGSWGGPPPSLAPRQEQAPGTOWPLDTSIGIYITIRERERESKGRPPWAAVVG 240
QY 241 PPOAHTSLQPHHHWPESVRESLCSLTPWKNEFNQKFTQLLLQRPHPRSQDPLVKRSW 300
Db 241 PPOAHTSLQPHHHWPESVRESLCSLTPWKNEFNQKFTQLLLQRPHPRSQDPLVKRSW 300
QY 301 PDYVEENRGHLIETRDIFGPGLDTOEPRIVILQGAAGICKSTARQVKEAWRGOLYGR 360
Db 301 PDYVEENRGHLIETRDIFGPGLDTOEPRIVILQGAAGICKSTARQVKEAWRGOLYGR 360
QY 361 FOHVFYFSCRELAQSKVSLAEIIGCKGTATPAPTRQILSRPERLLFLDGDVDEPGWLQ 420
Db 361 FOHVFYFSCRELAQSKVSLAEIIGCKGTATPAPTRQILSRPERLLFLDGDVDEPGWLQ 420
QY 421 EPSSELCLHWSQPQPADALLGSLGKLTILPEASFLITARTTALQNLIPSLQEARWVEVLG 480
```

Db 421 EFSSSELCHWSQPADALLGSLGKTLILPEASFLITARTTALQNLIPSELEQARWVEVLG 480
 Qy 481 FSESSRKEYFYFYFTDERQAIAPRLVKNKELWALCLVPWVSWLACTCLMOQKRKEKL 540
 Db 481 FSESSRKEYFYFYFTDERQAIAPRLVKNKELWALCLVPWVSWLACTCLMOQKRKEKL 540
 Qy 541 TLTSTKTTTTLCHYLAAQALQAPLGPOLRDLCSLAAEGIWQKTLFSPDDLRKHGLDGA 600
 Db 541 TLTSTKTTTTLCHYLAAQALQAPLGPOLRDLCSLAAEGIWQKTLFSPDDLRKHGLDGA 600
 Qy 601 ISTFLKMGILQHPHPIPLSYFIHLCPQEFPAAMSYYLEDEKRGKSHNSCIIIDLEKTXLEY 660
 Db 601 ISTFLKMGILQHPHPIPLSYFIHLCPQEFPAAMSYYLEDEKRGKSHNSCIIIDLEKTXLEY 660
 Qy 661 GIHLFGASTTFLGLLSDGEREMENIFHCRLSQGRNLMQWVPSLQLLQPHSLESILH 720
 Db 661 GIHLFGASTTFLGLLSDGEREMENIFHCRLSQGRNLMQWVPSLQLLQPHSLESILH 720
 Qy 721 CLYETRNTKTLQVMAHPEMGMCVETDMELLVCTFCIKFSRHVKQLQIEGRQHRSTWS 780
 Db 721 CLYETRNTKTLQVMAHPEMGMCVETDMELLVCTFCIKFSRHVKQLQIEGRQHRSTWS 780
 Qy 781 PTMVVLFWRVVPVTDAYWQILFVSLKVTRNLKELDLSGNSLSHSAVKSCLKTLRRPRCLLE 840
 Db 781 PTMVVLFWRVVPVTDAYWQILFVSLKVTRNLKELDLSGNSLSHSAVKSCLKTLRRPRCLLE 840
 Qy 841 TLRAGCGLTAECDKDLAFGLRANQTLTDLDSFNVLTDAGAKHLQRLQPSCKLQRLQ 900
 Db 841 TLRAGCGLTAECDKDLAFGLRANQTLTDLDSFNVLTDAGAKHLQRLQPSCKLQRLQ 900
 Qy 901 LVSCGLTSCCCDLASVLSASPSLKDLDLQNNLDDGVRLLCBGLRHPACKLIRGLDQ 960
 Db 901 LVSCGLTSCCCDLASVLSASPSLKDLDLQNNLDDGVRLLCBGLRHPACKLIRGLDQ 960
 Qy 961 TTLSDEMRQELRALEQEKQLLIFSRKPSVMTPTGGLDTGEMSNSTSLKQRQLGSE 1020
 Db 961 TTLSDEMRQELRALEQEKQLLIFSRKPSVMTPTGGLDTGEMSNSTSLKQRQLGSE 1020
 Qy 1021 ASHVAQANLKLDVSKIFPIAIEABESSPEVVPVELLCVPSASQGLHTKPLGTDGDFW 1080
 Db 1021 ASHVAQANLKLDVSKIFPIAIEABESSPEVVPVELLCVPSASQGLHTKPLGTDGDFW 1080
 Qy 1081 GPTGPVATEVDEKKNLYRHPVPVAGSYRWPNTGLCFVNRVATVEIEFCVWDQFLGEIN 1140
 Db 1081 GPTGPVATEVDEKKNLYRHPVPVAGSYRWPNTGLCFVNRVATVEIEFCVWDQFLGEIN 1140
 Qy 1141 PQHSMWVAGPLLDIAEPGAVEAVHLPHFVALQGGHVDTSLSFQMAHFKPEGMLLEKPARV 1200
 Db 1141 PQHSMWVAGPLLDIAEPGAVEAVHLPHFVALQGGHVDTSLSFQMAHFKPEGMLLEKPARV 1200
 Qy 1201 ELHHIVLENPSPGLGVLLKMHNLARFPTVTSVLLYHRVPEVTHLYLIPSDCSIR 1260
 Db 1201 ELHHIVLENPSPGLGVLLKMHNLARFPTVTSVLLYHRVPEVTHLYLIPSDCSIR 1260
 Qy 1261 KELELCYRSPGDDQLFSEFYVGHLSGIRLOVKDKDETLVWEALVKPGLMPATTLIPP 1320
 Db 1261 KELELCYRSPGDDQLFSEFYVGHLSGIRLOVKDKDETLVWEALVKPGLMPATTLIPP 1320
 Qy 1321 ARIAVPSPLDAPQLLHFVDYQREQLIARTVTSVEVVLKHLGQVLSQEQYERVLAENTRPS 1380
 Db 1321 ARIAVPSPLDAPQLLHFVDYQREQLIARTVTSVEVVLKHLGQVLSQEQYERVLAENTRPS 1380
 Qy 1381 QMRKLFSLSQSWDRCKDGLYQALKEPHLIMELWEKSKGLLPLSS 1429
 Db 1381 QMRKLFSLSQSWDRCKDGLYQALKEPHLIMELWEKSKGLLPLSS 1429

RESULT 5
 ABG71631
 ID ABG71631 standard; Protein; 1429 AA.
 XX
 AC ABG71631;

XX
 DT 09-JAN-2003 (first entry)
 XX
 DE Human caspase recruitment domain-7 (CARD-7).
 XX
 KW Human; caspase activity; caspase recruitment domain-7; CARD-7;
 KW caspase-1; pseudo-interleukin-1 beta converting enzyme; IL-1beta;
 KW pseudoICE; ICEBERG; cell growth; cell death; inflammation;
 KW apoptosis; caspase activation; cancer; follicular lymphoma;
 KW leukaemia; melanoma; colon cancer; lung carcinoma; viral infection;
 KW autoimmune disease; systemic lupus erythematosus; reactive arthritis;
 KW human immunodeficiency virus infection; HIV infection; AIDS;
 KW Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
 KW myocardial infarction; stroke; inflammatory disorder; Crohn's disease;
 KW insulin dependent diabetes mellitus; multiple sclerosis; psoriasis;
 KW graft rejection; allergic rhinitis; food allergy; conjunctivitis;
 KW glomerular nephritis; cytostatic; virucide; immunosuppressive;
 KW dermatological; nephrotropic; neuroprotective; cardiant.
 XX
 OS Homo sapiens.
 XX
 PN US2002128219-A1.
 XX
 PD 12-SEP-2002.
 XX
 PF 15-AUG-2001; 2001US-0931071.
 XX
 PR 27-OCT-1999; 99US-0428252.
 XX
 PA (BERT/) BERTIN J.
 PA (ALNE/) ALNEMRI E S.
 XX
 PI Bertin J, Alnemri ES;
 DR WPI; 2003-028968/02.
 DR N-PSDB; ABS55497.
 PT Assays for identifying compound that modulates the interaction of
 PT caspase recruitment domain-8 with a CARD-8 ligand or a compound that
 PT modulates activity of CARD-8 -
 PS Disclosure; Fig 1; 49pp; English.
 XX
 CC The present invention relates to methods of identifying compounds
 CC that regulate caspase activity using caspase recruitment domain-7
 CC (CARD-7) or caspase recruitment domain-8 (CARD-8). In particular, a
 CC method for identifying a compound that modulates the interaction
 CC between CARD-8 and caspase-1, pseudo-interleukin-1 (IL-1) beta
 CC converting enzyme (pseudoICE) or ICEBERG is disclosed. CARD-7 and
 CC CARD-8 molecules are useful as modulating agents in regulating a
 CC variety of cellular processes including cell growth, cell death, and
 CC inflammation. The methods of the invention are useful for identifying
 CC compounds that have the ability to increase/decrease apoptosis, or
 CC comprise the ability to induce caspase activation. The methods are
 CC useful for treating a disorder associated with inappropriate apoptosis
 CC or inappropriate inflammation. The methods are useful for treating
 CC disorders associated with an undesirably low rate of apoptosis such
 CC as cancer (preferably follicular lymphoma, chronic myelogenous
 CC leukaemia, melanoma, colon cancer, lung carcinoma, etc), viral
 CC infections, autoimmune diseases caused by low levels of apoptosis
 CC (e.g. systemic lupus erythematosus, immune-mediated glomerulonephritis,
 CC and arthritis). The methods are also useful for treating disorders with
 CC undesirably high rates of apoptosis such as human immunodeficiency
 CC virus (HIV) infection, Alzheimer's disease, Parkinson's disease,
 CC amyotrophic lateral sclerosis (ALS), retinitis pigmentosa, spinal
 CC muscular atrophy, various forms of cerebellar degeneration, anaemia
 CC associated with chronic disease, aplastic anaemia, chronic neutropaenia,
 CC myelodysplastic syndromes, myocardial infarction, stroke, and
 CC various inflammatory disorders (e.g. Crohn's disease, reactive
 CC arthritis, insulin dependent diabetes mellitus, multiple sclerosis,
 CC psoriasis, graft rejection, allergic rhinitis, food allergies,
 CC conjunctivitis, glomerular nephritis, etc). The present sequence
 CC represents human CARD-7.

XX	SQ	Sequence	1429 AA;
		Query Match 100.0%; Score 7534; DB 24; Length 1429;	
		Best Local Similarity 100.0%; Pred. No. 0;	
		Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1	MAGGAWGRLCAYLFBLKKEELKEFOLLANKAHSRSSGGETPAQPEKTSQGEVASYLVAQ	60
Db	1	MAGGAWGRLCAYLFBLKKEELKEFOLLANKAHSRSSGGETPAQPEKTSQGEVASYLVAQ	60
Qy	61	YGEORAWDLALHTWEQGLRSLCQAQEGAGHSFPFYPSPSEPHLGSPSOPSTAVLMPW	120
Db	61	YGEORAWDLALHTWEQGLRSLCQAQEGAGHSFPFYPSPSEPHLGSPSOPSTAVLMPW	120
Qy	121	IHELPACTQGSERRVLQRLPDTSGRRWREISALLYQALPSSPDHSPQESPNAPTST	180
Db	121	IHELPACTQGSERRVLQRLPDTSGRRWREISALLYQALPSSPDHSPQESPNAPTST	180
Qy	181	AVLGSWGSPPPSLAPREQEAFTQWPLDETSGIYYTEIREREREKSEKGRPPMAAVVGT	240
Db	181	AVLGSWGSPPPSLAPREQEAFTQWPLDETSGIYYTEIREREREKSEKGRPPMAAVVGT	240
Qy	241	PPQAHSTLOPHHHWPEVSRESLSTWPWKNEFNQKFTQLLLILORPHRPSQDPLVKRSW	300
Db	241	PPQAHSTLOPHHHWPEVSRESLSTWPWKNEFNQKFTQLLLILORPHRPSQDPLVKRSW	300
Qy	301	PDYVEENRGLHLEIRDLFGPGDLDTQEPRI VILQGAAGIGKSTLARQVKEAWGRGQLYDGR	360
Db	301	PDYVEENRGLHLEIRDLFGPGDLDTQEPRI VILQGAAGIGKSTLARQVKEAWGRGQLYDGR	360
Qy	361	FQHVYFYSCRELAQSKVVSIAELGKDGATAPAPIQLSRPERLLFILDGVEFGWVLQ	420
Db	361	FQHVYFYSCRELAQSKVVSIAELGKDGATAPAPIQLSRPERLLFILDGVEFGWVLQ	420
Qy	421	EPSELCLHWQPOPADALLGSLGKTLIPASFLITARTTALQNLPSLLEQARWVEVLG	480
Db	421	EPSELCLHWQPOPADALLGSLGKTLIPASFLITARTTALQNLPSLLEQARWVEVLG	480
Qy	481	FSESRKEYFYRYTFDERQAIAPRLVKSNEKELWALCLVPWVSWLACTCLMQQMKREKL	540
Db	481	FSESRKEYFYRYTFDERQAIAPRLVKSNEKELWALCLVPWVSWLACTCLMQQMKREKL	540
Qy	541	TLTSKTTTTLCHYLAAQALQAPLGPQLRDLCSLAAEGIWOKTKTFSPDDLKRHGLDGA	600
Db	541	TLTSKTTTTLCHYLAAQALQAPLGPQLRDLCSLAAEGIWOKTKTFSPDDLKRHGLDGA	600
Qy	601	ISTFLKMGILQEHPIPLYSYTHICFQOFFAAMSVYLEDEKRGKGNHCNIIDLEKTL EAY	660
Db	601	ISTFLKMGILQEHPIPLYSYTHICFQOFFAAMSVYLEDEKRGKGNHCNIIDLEKTL EAY	660
Qy	661	GIHGLFGASTTRFLGLLSDGGEREMENI FHCRLSQGNLMQWVPSLQLLQPHSLSLH	720
Db	661	GIHGLFGASTTRFLGLLSDGGEREMENI FHCRLSQGNLMQWVPSLQLLQPHSLSLH	720
Qy	721	CLYETRNKTFITQVMAHFEEMGMCVETDMLLVCTFCIKFSRHVKVQLQLEGRQHRSTWS	780
Db	721	CLYETRNKTFITQVMAHFEEMGMCVETDMLLVCTFCIKFSRHVKVQLQLEGRQHRSTWS	780
Qy	781	PTMVVLFRAWPVTDAYWQILFSLVKVTRNLKELDLSGNLSHSAVKSCKTLRRPCLLE	840
Db	781	PTMVVLFRAWPVTDAYWQILFSLVKVTRNLKELDLSGNLSHSAVKSCKTLRRPCLLE	840
Qy	841	TLRLACGGLTAEDCKDLAFLGRANQTLTELDLSFNVLTDAGAKHLQRLRQPSCKLQRLQ	900
Db	841	TLRLACGGLTAEDCKDLAFLGRANQTLTELDLSFNVLTDAGAKHLQRLRQPSCKLQRLQ	900
Qy	901	LVSCLGLTSDCCODLASVL SASPSL KELD LQNNLDDVGVRLLCEGLRHPACKLIRLGDQ	960
Db	901	LVSCLGLTSDCCODLASVL SASPSL KELD LQNNLDDVGVRLLCEGLRHPACKLIRLGDQ	960
Qy	961	TTLSDEMRQELRALEQEKPOLLI FSRRKPSVMTPTGELDTGEMNSNSTSSLKRQRLGSE	1020

AC AAY72669;
 XX 31-MAY-2001 (first entry)
 DE Human NB-ARC and CARD containing protein (NAC) beta isoform.
 DE Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;
 XX caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;
 KW cysteine aspartyl protease; apoptosis; cytokine production;
 KW cytokine receptor signalling; therapy; inflammatory disorder; sepsis;
 KW fibrosis; arthritis; cancer; adenocarcinoma; leukaemia.
 XX Homo sapiens.
 OS
 FH Key
 FH Domain
 FT Location/Qualifiers
 FT 329..547
 FT /label= NB domain
 FT /note= "Nucleotide binding domain, also designated as
 FT NB-ARC domain"
 FT 329..341
 FT /label= Walker A
 FT /note= "Also designated as P-loop"
 FT 406..414
 FT /label= Walker B
 FT 809..833
 FT /label= Leucine rich_repeat_region
 FT 838..862
 FT /label= Leucine rich_repeat_region
 FT 865..890
 FT /label= Leucine rich_repeat_region
 FT 895..919
 FT /label= Leucine rich_repeat_region
 FT 923..947
 FT /label= Leucine rich_repeat_region
 FT 957..987
 FT /note= "This 31 amino acid segment is not found in
 FT NAC gamma isoform (AAY72670) and NAC delta isoform
 FT (AAY72671) due to alternative mRNA splicing"
 FT 1261..1306
 FT /note= "This 45 amino acid segment is not found in
 FT NAC gamma isoform (AAY72670) due to alternative
 FT mRNA splicing"
 FT 1079..1364
 FT /note= "TIM-Barrel-like domain"
 FT 1128..1473
 FT /label= CARD-L
 FT /note= "Caspase-associated recruitment domain"
 FT 1128..1261
 FT /label= CARD-S
 FT /note= "Caspase-associated recruitment domain"
 FT 1306..1473
 FT /label= CARD-S
 FT /note= "Caspase-associated recruitment domain"
 FT 1373..1473
 FT /label= CARD
 FT /note= "Caspase-associated recruitment domain"
 FT
 XX WO200116170-A2.
 PN
 XX 08-MAR-2001.
 PD
 XX 01-SEP-2000; 2000WO-US24152.
 PP
 XX 01-SEP-1999; 99US-0388221.
 PR
 XX (BURN-) BURNHAM INST.
 PA
 XX Reed JC;
 PI
 XX WPI; 2001-183258/18.
 DR N-PSDB; AAD02760.
 DR
 XX Novel nucleic acid encoding NB-ARC and caspase associated recruitment

PT domains, used to produce polypeptides for screening for modulators of
 PT apoptosis -
 XX Claim 15; Page 133-137; 184pp; English.
 XX
 CC The present sequence is a human NB-ARC and CARD containing protein
 CC (NAC) beta isoform. NAC beta isoform represents the NAC splice
 CC variant in which both the splice regions are present in the translated
 CC polypeptide. NAC protein comprises a nucleotide binding (NB) domain
 CC (also referred as NB-ARC domain), a caspase-associated recruitment
 CC domain (CARD) and a TIM-Barrel-like domain. The caspases, cysteine
 CC aspartyl proteases, are principal effectors of apoptosis. CARD containing
 CC NAC proteins are used for screening modulators that modulates apoptosis,
 CC cytokine production, cytokine receptor signalling and other cellular
 CC processes. NAC can act as an immunogen for the production of polyclonal
 CC and monoclonal antibodies. It can also be used to diagnose and treat
 CC inflammatory disorders such as sepsis, fibrosis and arthritis and cancer
 CC pathologies such as adenocarcinomas and leukaemias. that shown as
 CC Note: This sequence is stated as being the same as the specification. However the
 CC SEQ ID NO:2 (AAY72711) in figure 1A of the specification. However the
 CC sequences differ at several positions.
 XX
 SQ Sequence 1473 AA;
 Query Match 99.6%; Score 7502; DB 22; Length 1473;
 Best Local Similarity 97.0%; Pred. No. 0;
 Matches 1429; Conservative 0; Mismatches 0; Indels 44; Gaps 1;
 QY 1 MAGGAWGLACYLEFLKKEELKEFQLLLANKAKHRSRSSGETPAQPEKTSGMGEVASYLVAQ 60
 DB 1 MAGGAWGLACYLEFLKKEELKEFQLLLANKAKHRSRSSGETPAQPEKTSGMGEVASYLVAQ 60
 QY 61 YGEQRANDLALHTWEQMGRLSLCAQAEAGHSPSPFPYSPSEPHLGSPSQSTSTAVLMPW 120
 DB 61 YGEQRANDLALHTWEQMGRLSLCAQAEAGHSPSPFPYSPSEPHLGSPSQSTSTAVLMPW 120
 QY 121 IHEIPAGTCQGSRRVRLQPLDPTSGRRWRREISALILYQALPSSPDHESPSQESNAPTST 180
 DB 121 IHEIPAGTCQGSRRVRLQPLDPTSGRRWRREISALILYQALPSSPDHESPSQESNAPTST 180
 QY 181 AVLGWSGPPPSLAPREQEPAGTQWPLDETSGIYYTIREREREKSEKGPMPAAVVG 240
 DB 181 AVLGWSGPPPSLAPREQEPAGTQWPLDETSGIYYTIREREREKSEKGPMPAAVVG 240
 QY 241 PQAHTSLQPHHPWEPSPVRESLSTWPKNEDFNQKFTQLLLQRPSPQDPLVKESW 300
 DB 241 PQAHTSLQPHHPWEPSPVRESLSTWPKNEDFNQKFTQLLLQRPSPQDPLVKESW 300
 QY 301 PDYVEENRGLHIEIRDLFGPGLDTPQPRIVILQGAAGIKSTLARQVKEAWRGOLYGD 360
 DB 301 PDYVEENRGLHIEIRDLFGPGLDTPQPRIVILQGAAGIKSTLARQVKEAWRGOLYGD 360
 QY 361 FQHVFPSCRELAQSKVSLAELIGKDGATATPAPITQILSRPERLLIFLDGVDPEGW 420
 DB 361 FQHVFPSCRELAQSKVSLAELIGKDGATATPAPITQILSRPERLLIFLDGVDPEGW 420
 QY 421 EPSSELCLHWSQPOPADALLGSLGKTLTPASFLITARTTALQNLIPSEARWVEVLG 480
 DB 421 EPSSELCLHWSQPOPADALLGSLGKTLTPASFLITARTTALQNLIPSEARWVEVLG 480
 QY 481 FSESRKEYFYRYFTDERQAIAPRLVKSNKELMALCLVPVWSWLACTCLMQMKRKEKL 540
 DB 481 FSESRKEYFYRYFTDERQAIAPRLVKSNKELMALCLVPVWSWLACTCLMQMKRKEKL 540
 QY 541 TLTSKTTTTLCHYLAQALQAPLQPLQDLCSLAAGIWKQKTLFSPDDLKHLGDGAI 600
 DB 541 TLTSKTTTTLCHYLAQALQAPLQPLQDLCSLAAGIWKQKTLFSPDDLKHLGDGAI 600
 QY 601 ISTFLKMGILQEHPIPLSYSTHLCFOFFAAMS VVLEDEKGRGKHSNCIDLEKTL 660
 DB 601 ISTFLKMGILQEHPIPLSYSTHLCFOFFAAMS VVLEDEKGRGKHSNCIDLEKTL 660
 QY 661 GIHGLFGASTTRFLGLGLLSDEGEREMENI FHCRLSQGRNLMQWVPSLQLLQPHSLES 720

Db 661 GIHGLFGASTTRFLGLLSDEGEREMENIFHCRLSQGRNLMQVPSLQLLQPHSLESILH 720
 Qy 721 CLYETRNKTLFQVMAHFEEMGCVETDMLLVCTFCIKFSRHVKQLQIEGRQRSTWS 780
 Db 721 CLYETRNKTLFQVMAHFEEMGCVETDMLLVCTFCIKFSRHVKQLQIEGRQRSTWS 780
 Qy 781 PTMVVLFVRVVPVTDAYWQILFSLVKVTRNLKELDLSGNSLSHSAVKSLSCKTLRRPRCLLE 840
 Db 781 PTMVVLFVRVVPVTDAYWQILFSLVKVTRNLKELDLSGNSLSHSAVKSLSCKTLRRPRCLLE 840
 Qy 841 TLRLAGCGLTAECDKDLAFGLRANQPLTDLDSFNVLTDAGAKHLQRLRQPSCKLQRLIQ 900
 Db 841 TLRLAGCGLTAECDKDLAFGLRANQPLTDLDSFNVLTDAGAKHLQRLRQPSCKLQRLIQ 900
 Qy 901 LVSCGLTSDCCDLASVLSASPSLKLDELQNNLDDVGVRLLCEGLRHHPACKLIRGLDQ 960
 Db 901 LVSCGLTSDCCDLASVLSASPSLKLDELQNNLDDVGVRLLCEGLRHHPACKLIRGLDQ 960
 Qy 961 TTLSDEMROELALEOEKPOLLIFFRRKPSVMTPTGLDTGEMSNSTSLKRLQSGSRA 1020
 Db 961 TTLSDEMROELALEOEKPOLLIFFRRKPSVMTPTGLDTGEMSNSTSLKRLQSGSRA 1020
 Qy 1021 ASHVAQANLKLDDVSKIFPIAIEAESSPEVVPVELLCVPSASQGLHTKPLGTDDDFW 1080
 Db 1021 ASHVAQANLKLDDVSKIFPIAIEAESSPEVVPVELLCVPSASQGLHTKPLGTDDDFW 1080
 Qy 1081 GTPGPVATEVDKERNLYRHPFVAGSYRWPNTGLCFVNRREAVTVEIEFCVWDQFLGEIN 1140
 Db 1081 GTPGPVATEVDKERNLYRHPFVAGSYRWPNTGLCFVNRREAVTVEIEFCVWDQFLGEIN 1140
 Qy 1141 PQHSMWVAGPLDDIKAEPCGAVEAVHPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
 Db 1141 PQHSMWVAGPLDDIKAEPCGAVEAVHPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
 Qy 1201 ELHHIVLENPSPGLVLLKMHNLRFIPVTSVLLYHRVHPPEVTFHLYLIPSDCSIR 1260
 Db 1201 ELHHIVLENPSPGLVLLKMHNLRFIPVTSVLLYHRVHPPEVTFHLYLIPSDCSIR 1260
 Qy 1261 -----KELELCVSPGEDQLF 1276
 Db 1261 KAIDLEMKFQVRIHKPPPLPLYNGCRYTVSGSGSGLMELPLKELELCVSPGEDQLF 1320
 Qy 1277 SEFYVGHLSGIRLQVKDKDETLYWEALVKPGDLMPTATLIPPARIAVPSPLDAPQLIH 1336
 Db 1321 SEFYVGHLSGIRLQVKDKDETLYWEALVKPGDLMPTATLIPPARIAVPSPLDAPQLIH 1380
 Qy 1337 FVDQYREQLIARVTSVEVVDLKLHGQVLSQEQYERVLAEINTRPSQMRKLFSLSQSWDRKC 1396
 Db 1381 FVDQYREQLIARVTSVEVVDLKLHGQVLSQEQYERVLAEINTRPSQMRKLFSLSQSWDRKC 1440
 Qy 1397 KDGLYQALKETHPHLIMELWEKSKGLPLSS 1429
 Db 1441 KDGLYQALKETHPHLIMELWEKSKGLPLSS 1473

RESULT 8

AA72711
 ID AA72711 standard; Protein; 1473 AA.

XX AC

XX AC

XX AC

XX 31-MAY-2001 (first entry)

XX DE Human NAC beta isoform, alternative version.

XX KW Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;
 KW caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;
 KW cysteine aspartyl protease; apoptosis; cytokine production;
 KW cytokine receptor signalling; therapy; inflammatory disorder; sepsis;
 KW fibrosis; arthritis; cancer; adenocarcinoma; leukaemia.
 XX OS Homo sapiens.

XX PH Key Location/Qualifiers
 FT Domain 329..547
 FT /label= NB domain
 FT /note= "Nucleotide binding domain, also designated as
 FT NB-ARC domain"
 FT 329..341
 FT /label= Walker A
 FT /note= "Also designated as P-loop"
 FT 406..414
 FT /label= Walker_B
 FT 809..833
 FT /label= Leucine_rich_repeat_region
 FT 838..862
 FT /label= Leucine_rich_repeat_region
 FT 865..890
 FT /label= Leucine_rich_repeat_region
 FT 895..919
 FT /label= Leucine_rich_repeat_region
 FT 923..947
 FT /label= Leucine_rich_repeat_region
 FT 957..987
 FT /note= "This 31 amino acid segment is not found in
 FT NAC gamma isoform (AA72670) and NAC delta isoform
 FT (AA72671) due to alternative mRNA splicing"
 FT 1261..1306
 FT /note= "This 45 amino acid segment is not found in
 FT NAC gamma isoform (AA72670) due to alternative
 FT mRNA splicing"
 FT 1079..1364
 FT /note= "TIM-Barrel-like domain"
 FT 1128..1473
 FT /label= CARD-L
 FT /note= "Caspase-associated recruitment domain"
 FT 1128..1261
 FT /label= CARD-S
 FT /note= "Caspase-associated recruitment domain"
 FT 1298..1305
 FT /note= "Encoded by GGGATGCTGGAATACTCCCAAG"
 FT 1306..1473
 FT /label= CARD-S
 FT /note= "Caspase-associated recruitment domain"
 FT 1373..1473
 FT /label= CARD
 FT /note= "Caspase-associated recruitment domain"
 FT WO200116170-A2.
 XX 08-MAR-2001.
 XX 01-SEP-2000; 2000WO-US24152.
 XX 01-SEP-1999; 99US-0388221.
 XX (BURN-) BURNHAM INST.
 XX Reed JC;
 XX WPI; 2001-183258/18.
 XX N-PSDB; RAD02760.
 XX Novel nucleic acid encoding NB-ARC and caspase associated recruitment
 XX domains, used to produce polypeptides for screening for modulators of
 XX apoptosis -
 XX Claim 15; Fig 1A; 184pp; English.
 XX The present sequence is a human NB-ARC and CARD containing protein
 XX (NAC) beta isoform, alternative version. NAC beta isoform represents the
 XX NAC splice variant in which both the splice regions are present in the
 XX translated polypeptide. NAC protein comprises a nucleotide binding (NB)
 XX domain (also referred as NB-ARC domain), a caspase-associated recruitment
 XX domain (CARD) and a TIM-Barrel-like domain. The caspases, cysteine

aspartyl proteases, are principal effectors of apoptosis. CARD containing NAC proteins are used for screening modulators that modulate apoptosis, cytokine production, cytokine receptor signalling and other cellular processes. NAC can act as an immunogen for the production of polyclonal and monoclonal antibodies. It can also be used to diagnose and treat inflammatory disorders such as sepsis, fibrosis and arthritis and cancer pathologies such as adenocarcinomas and leukaemias.
 Note: This sequence is stated as being the same as that shown as SEQ ID NO:2 (See AAY72669) in page 133-137 of the specification. However the sequences differ at several positions.

CC	Query Match	99.6%;	Score 7502;	DB 22;	Length 1473;
CC	Best Local Similarity	97.0%;	Pred. No. 0;		
CC	Matches 1429;	Conservative 0;	Mismatches 0;	Indels 44;	Gaps 1;
QY	1	MAGGAWRLACVLEFLKKEELKEFQALLANKAHSRSSSGETPAQPEKTSGMVEASVLYNAQ	60		
DB	1	MAGGAWRLACVLEFLKKEELKEFQALLANKAHSRSSSGETPAQPEKTSGMVEASVLYNAQ	60		
QY	61	YGEQRAWDALHTWEOQLRSICAQAEAGAGHSPPYSPSPHLCSPSQPTSTAVLMPW	120		
DB	61	YGEQRAWDALHTWEOQLRSICAQAEAGAGHSPPYSPSPHLCSPSQPTSTAVLMPW	120		
QY	121	IHELPACTQGSERRVLRQLPDTSGRRWREISASLLYQALPSPDHESPSQSPNAPTST	180		
DB	121	IHELPACTQGSERRVLRQLPDTSGRRWREISASLLYQALPSPDHESPSQSPNAPTST	180		
QY	181	AVLGSWSPQPSLAPREOAGTQWPLDETSGIYYTEIREREREKSEKGRPPAAVGT	240		
DB	181	AVLGSWSPQPSLAPREOAGTQWPLDETSGIYYTEIREREREKSEKGRPPAAVGT	240		
QY	241	PQOAHSLQPHHPWPSPRESICSTWPKWNEFNOKFTQLLLQPHRPSQDPLVRSW	300		
DB	241	PQOAHSLQPHHPWPSPRESICSTWPKWNEFNOKFTQLLLQPHRPSQDPLVRSW	300		
QY	301	PDYVENRGLHIEIRDLFGPLDTPRIVILQGAAGIKSTLARQVKEAWGRGQLYGR	360		
DB	301	PDYVENRGLHIEIRDLFGPLDTPRIVILQGAAGIKSTLARQVKEAWGRGQLYGR	360		
QY	361	FQHVFFSCRELAQSVLSAEILGDKGTATPAPIRQILSRPERLLFILDGVDPEGWLQ	420		
DB	361	FQHVFFSCRELAQSVLSAEILGDKGTATPAPIRQILSRPERLLFILDGVDPEGWLQ	420		
QY	421	EPSELCLHWSQPADALLGSLGKTLILPEASFLITARTALQNLIPSLQARWVEVLG	480		
DB	421	EPSELCLHWSQPADALLGSLGKTLILPEASFLITARTALQNLIPSLQARWVEVLG	480		
QY	481	FSESSRKEYFYRTDQRAIRAFRLVKNKELWALCLVPWVSWLACTCLMQMKRKEKL	540		
DB	481	FSESSRKEYFYRTDQRAIRAFRLVKNKELWALCLVPWVSWLACTCLMQMKRKEKL	540		
QY	541	TLTSKTTTTLCHYLAAQALQAPLQPLRDCLSLAAEGIWQKTLFSPDRLRKHGLDGA	600		
DB	541	TLTSKTTTTLCHYLAAQALQAPLQPLRDCLSLAAEGIWQKTLFSPDRLRKHGLDGA	600		
QY	601	ISTFLKMGILQEHPIPLSYFHLCPQEPFAAMSVYLEDEKGRKHSNCIILDEKLTLEY	660		
DB	601	ISTFLKMGILQEHPIPLSYFHLCPQEPFAAMSVYLEDEKGRKHSNCIILDEKLTLEY	660		
QY	661	GIHGLFGASTTRFLGLLSDGEREMENIFCHRLSQGRNLQMWVPSLQLLQPHSLH	720		
DB	661	GIHGLFGASTTRFLGLLSDGEREMENIFCHRLSQGRNLQMWVPSLQLLQPHSLH	720		
QY	721	CLYETRNKTLTQVMAHFEBMGVCVETDMLLYCTFCIKFSRHVKKLQIEGRQHSWTS	780		
DB	721	CLYETRNKTLTQVMAHFEBMGVCVETDMLLYCTFCIKFSRHVKKLQIEGRQHSWTS	780		
QY	781	PTWVLFVRVPTVDAYWQILFSLVKVTRNLKELDLGNSLSHSAVSKLCTLRPRCLLE	840		
DB	781	PTWVLFVRVPTVDAYWQILFSLVKVTRNLKELDLGNSLSHSAVSKLCTLRPRCLLE	840		

QY	841	TURLAGCGLTABDCXDLAFGLRANQTLTDLDFSNVLTDAKAKHLCOBLRQPSCKLRLQ	900		
DB	841	TURLAGCGLTABDCXDLAFGLRANQTLTDLDFSNVLTDAKAKHLCOBLRQPSCKLRLQ	900		
QY	901	LVSCLGTSDDCCDLASVLSASPSLKELDLQNNLDDVGRLLCEGLRHHPACKLIIRGLDQ	960		
DB	901	LVSCLGTSDDCCDLASVLSASPSLKELDLQNNLDDVGRLLCEGLRHHPACKLIIRGLDQ	960		
QY	961	TTLSDEMRQELRALBQEKPOLLIIFSRKPSVMTPTTEGLDTGBMSNSTSLKQRKLSERA	1020		
DB	961	TTLSDEMRQELRALBQEKPOLLIIFSRKPSVMTPTTEGLDTGBMSNSTSLKQRKLSERA	1020		
QY	1021	ASHVAQANLLDVSKEIPFIAIABESSPEVVPVELLCVPSPASQGLHTRKPLGTTDDFW	1080		
DB	1021	ASHVAQANLLDVSKEIPFIAIABESSPEVVPVELLCVPSPASQGLHTRKPLGTTDDFW	1080		
QY	1081	GTPGVATEVVDKKNLYRVHFPVAGSYRWPNTGLCFVNRVAVTVEIEFCVWDQFLGEIN	1140		
DB	1081	GTPGVATEVVDKKNLYRVHFPVAGSYRWPNTGLCFVNRVAVTVEIEFCVWDQFLGEIN	1140		
QY	1141	PQHSMMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV	1200		
DB	1141	PQHSMMVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV	1200		
QY	1201	ELHHIVLENPSPLGVLLKMIHNAIRFIPVTSVLLVHRVHPPEVTHLYLIPSDCSIR	1260		
DB	1201	ELHHIVLENPSPLGVLLKMIHNAIRFIPVTSVLLVHRVHPPEVTHLYLIPSDCSIR	1260		
QY	1261	K-----EELCYRSPGEDQLF	1276		
DB	1261	KAIDDLWKPFQVRHKKPPPLTPLYMGCRYTVSSGSRDAGNTPOELELCYRSPGEDQLF	1320		
QY	1277	SEFYVGHGSGIRLOVKDKOBTLLVWEALVKGDLMPATTLIPPARIAVPSPLDAPQLLH	1336		
DB	1321	SEFYVGHGSGIRLOVKDKOBTLLVWEALVKGDLMPATTLIPPARIAVPSPLDAPQLLH	1380		
QY	1337	FVDQYREQLIARVTSVEVVDKLHGQVLSQEQYERVALENTRPSQMRKLFSLSQSWDRKC	1396		
DB	1381	FVDQYREQLIARVTSVEVVDKLHGQVLSQEQYERVALENTRPSQMRKLFSLSQSWDRKC	1440		
QY	1397	KDGLYQALKEHPLHLMELWEKSKKGLPLSS	1429		
DB	1441	KDGLYQALKEHPLHLMELWEKSKKGLPLSS	1473		
RESULT 9					
ABG78455					
ID	ABG78455	standard; Protein; 1429 AA.			
XX	AC	ABG78455;			
XX	DT	15-NOV-2002 (first entry)			
XX	DE	Human caspase recruitment protein 7 protein.			
XX	KW	Human; human leucine-rich repeat small intestine I; HLRRSII; asthma;			
XX	KW	proliferative disorder; gastrointestinal disorder; renal disorder;			
XX	KW	neural disorder; reproductive disorder; calcium regulation; apoptosis;			
XX	KW	immune system; anaemia; human immune deficiency virus; HIV; cancer;			
XX	KW	blood coagulation disorder; autoimmune disorder; allergic reaction;			
XX	KW	inflammatory condition; cardiovascular disorder; ischaemia;			
XX	KW	neurological disorder; infectious disease; cytokine production;			
XX	OS	expressed sequence tag; EST.			
XX	PN	Homo sapiens.			
XX	PN	WO200261086-A2.			
XX	PD	08-AUG-2002.			
XX	PF	20-DEC-2001; 2001WO-US49739.			
XX	PR	22-DEC-2000; 2000US-257774P.			

XX	20-AUG-2002	(first entry)	QY	301	PDYVEENRGLHIEIRDLFGPGLDTPQPRIVILQGAAGIKSTLARQVKEAWRGQLYGDR	360
DT			Db	301	PDYVEENRGLHIEIRDLFGPGLDTPQPRIVILQGAAGIKSTLARQVKEAWRGQLYGDR	360
DE						
XX		Pyrin domain containing protein NALP1-hs.	QY	361	FOHVYFSCRELAQSKVSLAELICKDGTATPAPIRQILSRPERLLFILDGVDPEGWVQ	420
KW		Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;	Db	361	FOHVYFSCRELAQSKVSLAELICKDGTATPAPIRQILSRPERLLFILDGVDPEGWVQ	420
KW		antiartherosclerotic; anticypriotic; antibacterial; virucide;				
KW		neuroprotective; antiarthritic; antirheumatic; antiasthmatic;	QY	421	EPSSSELCHWSQPADALLGSLGKTILPEASFLITARTTALQNLIPSELEQARVVEVLG	480
KW		nephrotropic; osteopathic; nootropic; intracellular signal transduction;	Db	421	EPSSSELCHWSQPADALLGSLGKTILPEASFLITARTTALQNLIPSELEQARVVEVLG	480
KW		inflammation; Alzheimer's disease; infection; psoriasis; asthma;				
KW		arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;	QY	481	PSSESRKEYFYRYFTDERQATRAFRVLSKNELWALCLVPVWVSWLACTCLMOMKRRKEKL	540
KW		osteoarthritis; glomerulonephritis.	Db	481	PSSESRKEYFYRYFTDERQATRAFRVLSKNELWALCLVPVWVSWLACTCLMOMKRRKEKL	540
OS		Unidentified.				
XX		WO200240668-A2.	QY	541	TLTSTKTTTTLCHYLAQALQAPLGPQRLDCLSLAAGIWOKKTLFSPDDLRKHGLDGA	600
PN			Db	541	TLTSTKTTTTLCHYLAQALQAPLGPQRLDCLSLAAGIWOKKTLFSPDDLRKHGLDGA	600
XX		23-MAY-2002.				
XX		30-OCT-2001; 2001WO-BF12545.	QY	601	ISTFLKMGILQEHPIPLSYSFHLCQBFFAAMS YVLEDEKGRGKHSNCIIDLKTLKAY	660
PF			Db	601	ISTFLKMGILQEHPIPLSYSFHLCQBFFAAMS YVLEDEKGRGKHSNCIIDLKTLKAY	660
XX		15-NOV-2000; 2000DE-1056687.				
PR			QY	661	GIHGLFGASTTRFLGLLSDEGEREMENI FHCRLSQGRNLMQWVPSLQLLQPHSLESILH	720
PR		30-NOV-2000; 2000DE-1059595.	Db	661	GIHGLFGASTTRFLGLLSDEGEREMENI FHCRLSQGRNLMQWVPSLQLLQPHSLESILH	720
XX						
XX		(APOT-) APOTECH RES & DEV LTD.	QY	721	CLYETRNKTFLTQVMAHFEEMCMVETDMLLVCTFCIKFSRHVKKLIQIEGRQHSRWS	780
PA			Db	721	CLYETRNKTFLTQVMAHFEEMCMVETDMLLVCTFCIKFSRHVKKLIQIEGRQHSRWS	780
XX		Tschopp J, Martinon F;				
PI		WPI: 2002-427093/45.	QY	781	PTMVVLFWRVPTDAYWQILFSLVLTNRNLKELDLGNSLSHSAVSKLCTLRPRCLLE	840
XX		N-PSDB; AAL47127.	Db	781	PTMVVLFWRVPTDAYWQILFSLVLTNRNLKELDLGNSLSHSAVSKLCTLRPRCLLE	840
XX						
PT		New DNA encoding protein with pyrin domain, useful for treating	QY	841	TLRLAGCGLTABDCXDLAFGLRANQTLTDLDFSNVLTDAKAKHLCQRLRQPSCKLQRLQ	900
PT		diseases involving impaired signal transduction, particularly	Db	841	TLRLAGCGLTABDCXDLAFGLRANQTLTDLDFSNVLTDAKAKHLCQRLRQPSCKLQRLQ	900
PT		inflammation, also proteins and antibodies				
XX		Claim 5; Fig 1; 116pp; German.	QY	901	LVS CGLTSDCCODLASVLSASPSLKLQNNLDVGVLLCEGLRHACKLIRGLDQ	960
XX		The present invention relates the DNA and their encoded proteins, where	Db	901	LVS CGLTSDCCODLASVLSASPSLKLQNNLDVGVLLCEGLRHACKLIRGLDQ	960
CC		the proteins contain at least one PYD (pyrin) domain. These can be used				
CC		to treat diseases associated with impaired intracellular signal	QY	961	TTLSDEMQRALRQEKPOLLI FRRKPSVMTPTGDLTGEMSNSTSLKQRKLSERA	1020
CC		transduction, particularly inflammation such as psoriasis,	Db	961	TTLSDEMQRALRQEKPOLLI FRRKPSVMTPTGDLTGEMSNSTSLKQRKLSERA	1020
CC		arteriosclerosis, bacterial or viral infections (particularly meningitis				
CC		and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,	QY	1021	ASHVAQANLKLDVSKIPPIAIEAESSESPVVPVELLCVPSASQGLHTKPLGTDGDFW	1080
CC		sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's	Db	1021	ASHVAQANLKLDVSKIPPIAIEAESSESPVVPVELLCVPSASQGLHTKPLGTDGDFW	1080
CC		and Parkinson's diseases. The present sequence is a protein of the				
CC		invention.	QY	1081	GTPGVATVVDKCNLYRVHPFVAGSVRWPNGLCFVWREAVTVIEIFCVWDQFLGEIN	1140
XX			Db	1081	GTPGVATVVDKCNLYRVHPFVAGSVRWPNGLCFVWREAVTVIEIFCVWDQFLGEIN	1140
SQ						
		Query Match 99.4%; Score 7488; DB 23; Length 1429;				
		Best Local Similarity 99.4%; Pred. No. 0;				
		Matches 1420; Conservative 5; Mismatches 4; Indels 0; Gaps 0;				
QY	1	MAGGAWGRACYLEFLFKXGELKEFQLLANKAHSRSSGETPAQPKTSGMEVASVYLAQ	60			
Db	1	MAGGAWGRACYLEFLFKXGELKEFQLLANKAHSRSSGETPAQPKTSGMEVASVYLAQ	60			
QY	61	YGEQRAWDLALHTWEQGLRSLCAQAEAGHSPSPYSPSEPHLGSQPTSTAVLMPW	120			
Db	61	YGEQRAWDLALHTWEQGLRSLCAQAEAGHSPSPYSPSEPHLGSQPTSTAVLMPW	120			
QY	121	IHELPACTQGSRRVRLQPLDTSGRWRREISASLLYQALPSPDHSPSQSPNAPTST	180			
Db	121	IHELPACTQGSRRVRLQPLDTSGRWRREISASLLYQALPSPDHSPSQSPNAPTST	180			
QY	181	AVLWGSQPPQPSLAPREQEAPGTQWPLDTSYIYTEREREREKSRPPAAVVGCT	240			
Db	181	AVLWGSQPPQPSLAPREQEAPGTQWPLDTSYIYTEREREREKSRPPAAVVGCT	240			
QY	241	PPQAHSLQPHHPWPSPVRESLCSSTWPKNEDFNQKFTQLLLQRPSPRQDPLVKRSW	300			
Db	241	PPQAHSLQPHHPWPSPVRESLCSSTWPKNEDFNQKFTQLLLQRPSPRQDPLVKRSW	300			

Db	1381	QMRKLFSLQSWDRKCKDGLYQALKETHPHILMELWEKSKKGLPLSS	1429
RESULT 11			
AAE06758			
ID	AAE06758	standard; Protein; 1473 AA.	
AC	AAE06758;		
XX			
XX	16-OCT-2001	(first entry)	
DE	Human	G-protein coupled receptor-8 (GCRC-8) protein.	
XX	Human:	G-protein coupled receptor-8; GCRC-8; cytostatic; hepatotropic; virucide; antiinflammatory; anticonvulsant; antiemetic; neuroprotective; nootropic; cerebroprotective; hypotensive; tranquiliser; vulnerary; ophthalmological; cell proliferative disorder; actinic keratosis; anorectic; arteriosclerosis; atherosclerosis; cirrhosis; hepatitis; psoriasis; cancer; neurological disorder; stroke; Alzheimer's disease; Huntington's disease; Parkinson's disease; cardiovascular disorder; epilepsy; hypertension; varicose vein; vasculitis; dysphagia; dyspepsia; anorexia; gastrointestinal disorder; pancreatitis; autoimmune disease; Addison's disease; Crohn's disease; acquired immune deficiency syndrome; AIDS; uveitis; infection; trauma; metabolic disorder; diabetes; obesity; osteoporosis; transgenic animal; gene therapy.	
OS	Homo sapiens.		
XX	Key	Location/Qualifiers	
FT	Domain	1216..1237	
FT		/label= Transmembrane_domain	
FT	Binding-site	334..341	
FT		/label= ATP/GTP-binding_site	
FT		/note= "P-loop"	
XX	WO200157085-A2.		
PN	09-AUG-2001.		
XX	01-FEB-2001;	2001WO-US03455.	
XX	02-FEB-2000;	2000US-0180093.	
PR	11-FEB-2000;	2000US-0182045.	
XX	(INCY-)	INCYTE GENOMICS INC.	
PA	Baughn MR,	Au-Young J, Yue H;	
PI	WPI: 2001-488869/53.		
DR	N-PSDB; AAD12951.		
XX	Novel isolated human G-protein coupled receptor useful for diagnosing, preventing and treating cell proliferative, neurological, cardiovascular, gastrointestinal, autoimmune/inflammatory and metabolic disorders -		
PS	Claim 1;	Page 114-117; 138pp; English.	
XX	The present sequence is human G-protein coupled receptor-8 (GCRC-8) protein. The present invention relates to GCRC protein and nucleic acids encoding them. GCRC protein, its agonist or antagonist are useful for treating diseases or conditions associated with decreased expression or overexpression of functional GCRC in a patient, where the disorder is selected from cell proliferative disorders such as actinic keratosis, arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, psoriasis, and cancer, neurological disorders such as epilepsy, stroke, Alzheimer's disease, Huntington's disease, Parkinson's disease, cardiovascular disorders such as hypertension, vasculitis, varicose veins, gastro-intestinal disorders such as dysphagia, dyspepsia, anorexia, nausea, pancreatitis, autoimmune/inflammatory disorders such as acquired immunodeficiency syndrome (AIDS), Addison's disease, Crohn's disease, uveitis, viral, bacterial, fungal, parasitic, protozoal, helminthic		

CC infections, trauma and metabolic disorders such as diabetes, obesity, CC osteoporosis. GCRC proteins and their cDNAs are used to assess the CC effects of exogenous compounds on the expression of GCRC sequences. CC GCRC cDNA is useful to create knock in humanised animals (pigs) or CC transgenic animals (mice or rats) to model human disease, for CC therapeutic or diagnostic purposes, for somatic or germline gene CC therapy, to generate hybridisation probes useful in mapping the CC naturally occurring genomic sequence, and in molecular biological CC techniques.

XX	Sequence	1473 AA;	
XX	Query Match	99.3%; Score 7484; DB 22; Length 1473;	
XX	Best Local Similarity	96.8%; Pred. No. 0;	
XX	Matches 1426;	Conservative 1; Mismatches 2; Indels 44; Gaps 1;	
QY	1	MAGGAWGLACYLEFLKKEELKEFOLLANKAHSSSGGETPAQPEKTSQMEVASYLVAQ	60
DB	1	MAGGAWGLACYLEFLKKEELKEFOLLANKAHSSSGGETPAQPEKTSQMEVASYLVAQ	60
QY	61	YGEQAWDLALHTWQMGRLSLCAQAEAGAGHSFPSPSEPHLGSPQSTSTAVLMPW	120
DB	61	YGEQAWDLALHTWQMGRLSLCAQAEAGAGHSFPSPSEPHLGSPQSTSTAVLMPW	120
QY	121	IHELPAQCTQGSRRVLRQLPDTSGRRWRREISASLLYQALPSSPDHESPQSPNAPTST	180
DB	121	IHELPAQCTQGSRRVLRQLPDTSGRRWRREISASLLYQALPSSPDHESPQSPNAPTST	180
QY	181	AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTIREREREKSEKGRPPMAVVG	240
DB	181	AVLGSWGSPPQPSLAPREQEAPGTQWPLDETSGIYYTIREREREKSEKGRPPMAVVG	240
QY	241	PPQAHTSLQPHHPWEPSPVRESLCSLTPWQKEDFNQKTQLLLQRPSPQDPLVKESW	300
DB	241	PPQAHTSLQPHHPWEPSPVRESLCSLTPWQKEDFNQKTQLLLQRPSPQDPLVKESW	300
QY	301	PDYVEENRGLHIEIRDLFGPLDTPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGDR	360
DB	301	PDYVEENRGLHIEIRDLFGPLDTPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGDR	360
QY	361	FQHFVYFSCRELAQSKVVSALBLIGKDGATATPAPIRQILSRPERLLFLDGVDEPGWLQ	420
DB	361	FQHFVYFSCRELAQSKVVSALBLIGKDGATATPAPIRQILSRPERLLFLDGVDEPGWLQ	420
QY	421	EPSELCLHWSQPADALLGSLGKTLTPASFLITARTTALQNLIPSEQARVVEVLG	480
DB	421	EPSELCLHWSQPADALLGSLGKTLTPASFLITARTTALQNLIPSEQARVVEVLG	480
QY	481	FSESRKEYFYRYFTDERQAIAPRLVKSNEKELWALCLVPWVSWLACTCLMQMKRKEKL	540
DB	481	FSESRKEYFYRYFTDERQAIAPRLVKSNEKELWALCLVPWVSWLACTCLMQMKRKEKL	540
QY	541	TLTSKTTTTLCLHLYLAQALQAPLGPQLRDLCSLAAGIWKQKTLFSPDDLRKHGLDGA	600
DB	541	TLTSKTTTTLCLHLYLAQALQAPLGPQLRDLCSLAAGIWKQKTLFSPDDLRKHGLDGA	600
QY	601	ISTFLKMGILQEHPIPLSYFIHLFCQFFPAMSVVLEDEKRGKHSNCIIDLEKTLRAY	660
DB	601	ISTFLKMGILQEHPIPLSYFIHLFCQFFPAMSVVLEDEKRGKHSNCIIDLEKTLRAY	660
QY	661	GIHGLFGASTTRFLGLLSDSGEREMENIFHCRLSQGRNLQMWVPSLQLLQPHSLES	720
DB	661	GIHGLFGASTTRFLGLLSDSGEREMENIFHCRLSQGRNLQMWVPSLQLLQPHSLES	720
QY	721	CLYETRNKTLTQVMAHFEEMGCMVETDMLLVCTFCIKFSRHVKKLQLEIGRQHRSTWS	780
DB	721	CLYETRNKTLTQVMAHFEEMGCMVETDMLLVCTFCIKFSRHVKKLQLEIGRQHRSTWS	780
QY	781	PTMVVLFWRVPTDAYWQILSVLKVTRNLKELDLSGNSHSAVKSICKTLRRRCILLE	840
DB	781	PTMVVLFWRVPTDAYWQILSVLKVTRNLKELDLSGNSHSAVKSICKTLRRRCILLE	840
QY	841	TLRLAGCGLTAEDCKDLAFGLRANQTLTDLSPFNVLTDAGAKHLQRLRQPSCKLQRLQ	900

1 MAGGAWRLACYLEFFLKKKEELKEFQLLANKAHRSSSGETPAQPEKTSMEVASYLVAQ
1 MAGGAWRLACYLEFFLKKKEELKEFQLLANKAHRSSSGETPAQPEKTSMEVASYLVAQ

DT 31-MAY-2001 (first entry)

Db 181 AVLSGWSPPQSLAPREQAPGTQWPLDETSGIYYTEIRERERESKGRPPMAVVTG 240
Qy 241 PQAHTSLQPHHPHPEPSVRESLCSWTWPKNEDFNQKFTQLLLQRPHPRSODPLVKRSW 300
Db 241 PQAHTSLQPHHPHPEPSVRESLCSWTWPKNEDFNQKFTQLLLQRPHPRSODPLVKRSW 300
Qy 301 PDYVENRGLHLEIRDLFGFGLDTPRIVILQGAAGIKSTLARQVKEAMGRGQLYGDR 360
Db 301 PDYVENRGLHLEIRDLFGFGLDTPRIVILQGAAGIKSTLARQVKEAMGRGQLYGDR 360
Qy 361 FOHVYFSCRELAQSKVSLAELIKGCGTATPAPITRQILSRPERLLFILDGDEGFWLQ 420
Db 361 FOHVYFSCRELAQSKVSLAELIKGCGTATPAPITRQILSRPERLLFILDGDEGFWLQ 420
Qy 421 EFSSELCHWSQPADALLGSLGKTILPEASFLITARTTALQNLIPSLQARWVEVLG 480
Db 421 EFSSELCHWSQPADALLGSLGKTILPEASFLITARTTALQNLIPSLQARWVEVLG 480
Qy 481 FSESRKEYFYRYFTDERQAIRAFRLVKSNNKELWALCLVPWVSWLACTCLMQMKRKEKL 540
Db 481 FSESRKEYFYRYFTDERQAIRAFRLVKSNNKELWALCLVPWVSWLACTCLMQMKRKEKL 540
Qy 541 TLTSKTTTTLCHYLALQALQAPGLQPLDLSLAAEGIWQKTLFSPDDLRKHGLDGA 600
Db 541 TLTSKTTTTLCHYLALQALQAPGLQPLDLSLAAEGIWQKTLFSPDDLRKHGLDGA 600
Qy 601 ISTFLKMGILQHPHPLPSYSFTHLQFQEFFAAMSVYLEDKGRGKHSNCIIIDLEKTLAY 660
Db 601 ISTFLKMGILQHPHPLPSYSFTHLQFQEFFAAMSVYLEDKGRGKHSNCIIIDLEKTLAY 660
Qy 661 GIHGLFGASTTRFLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLQPHLSLSLH 720
Db 661 GIHGLFGASTTRFLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLQPHLSLSLH 720
Qy 721 CLYETRNKTLFQVWAHFEEMGCVETDMELLVCTFCIKFSRHVKQLQIQRQHRSTWS 780
Db 721 CLYETRNKTLFQVWAHFEEMGCVETDMELLVCTFCIKFSRHVKQLQIQRQHRSTWS 780
Qy 781 PTMVVLFRVWPVTDAWQILFVSLKVRNLKELDLSGNSLSHSAVSKLCKTLRRPRCLLE 840
Db 781 PTMVVLFRVWPVTDAWQILFVSLKVRNLKELDLSGNSLSHSAVSKLCKTLRRPRCLLE 840
Qy 841 TLRLACGLTAEDCKDLAFLGRANQTLTDLDSFNVLTDAGAKHLQRLRQPSCKLQRLQ 900
Db 841 TLRLACGLTAEDCKDLAFLGRANQTLTDLDSFNVLTDAGAKHLQRLRQPSCKLQRLQ 900
Qy 901 LVSCGLTSCCQDLASVLSASPSLKELDLQNNLDDGVVRLLCGLRHHPACKLIRGLDQ 960
Db 901 LVSCGLTSCCQDLASVLSASPSLKELDLQNNLDDGVVRLLCGLRHHPACKLIRL 956
Qy 961 TTLSDEMRLQALQEKQQLLI FSRKPSVMTPTGLDTGEMSNSTSLKQRLGSERA 1020
Db 957 -----KPSVMTPTGLDTGEMSNSTSLKQRLGSERA 989
Qy 1021 ASHVAQANLKLDVSKIFPIAETAEBSSEVVPVELLCVPSASQGLDHTKPLGTDGDFW 1080
Db 990 ASHVAQANLKLDVSKIFPIAETAEBSSEVVPVELLCVPSASQGLDHTKPLGTDGDFW 1049
Qy 1081 GPTGPVATEVDEKKNLYRVHPFPVAGSYRWPNTGLCFVNRVATVIEFCVNDQFLGEIN 1140
Db 1050 GPTGPVATEVDEKKNLYRVHPFPVAGSYRWPNTGLCFVNRVATVIEFCVNDQFLGEIN 1109
Qy 1141 PQHSNNVAGLLDIAKEPQAVEAVHLPHFVALQGGHVDTSLFQMAHFEKGMLEKPARV 1200
Db 1110 PQHSNNVAGLLDIAKEPQAVEAVHLPHFVALQGGHVDTSLFQMAHFEKGMLEKPARV 1169
Qy 1201 ELHHIVLENPSPGLGVLLKMHNAURLFIPVTSVLLYHRVPEEVTFFHLYLIPSDCSR 1260
Db 1170 ELHHIVLENPSPGLGVLLKMHNAURLFIPVTSVLLYHRVPEEVTFFHLYLIPSDCSR 1229
Qy 1261 -----KELELCYRSPGSDQFL 1276
Db 1230 KATDDLEMKFQFVR IHKPPPLTPLYMCRYTVYSGSGMGLTILPKELELCYRSPGSDQFL 1289

Qy 1277 SEFYVGHLSGIRLQVKDKKDETLVWEALVKPGDLMPATTLIPPARIAVPSPLDAPQLLH 1336
Db 1290 SEFYVGHLSGIRLQVKDKKDETLVWEALVKPGDLMPATTLIPPARIAVPSPLDAPQLLH 1349
Qy 1337 FVDQYREQLIARVTSVEVLDKLHGQVLSQBOYERVAENTRPSQMRKLPSSLQSWDRKC 1396
Db 1350 FVDQYREQLIARVTSVEVLDKLHGQVLSQBOYERVAENTRPSQMRKLPSSLQSWDRKC 1409
Qy 1397 KDGLYQALKETHPHLIMELWEKSKKGLLPLSS 1429
Db 1410 KDGLYQALKETHPHLIMELWEKSKKGLLPLSS 1442
RESULT 14
AAV72673
ID AAV72673 standard; Protein; 1454 AA.
XX
AC AAV72673;
XX
DT 31-MAY-2001 (first entry)
XX
DE Human NAC beta isoform-CARD-X1 chimeric protein.
XX
KW Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;
KW caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;
KW cysteine aspartyl protease; apoptosis; cytokine production;
KW cytokine receptor signalling; therapy; inflammatory disorder; sepsis;
KW fibrosis; arthritis; cancer; adenocarcinoma; leukaemia;
KW chimeric protein.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 1..1078
FT /note= "Corresponds to 1-1078 residues of human NAC
FT beta isoform (AAV72669)"
FT Region 1079..1454
FT /note= "Corresponds to 56-431 residues of human CARD-X
FT protein (AAV72672)"
XX WO200116170-A2.
XX
PD 08-MAR-2001.
XX
PF 01-SEP-2000; 2000WO-US24152.
XX
PR 01-SEP-1999; 99US-0388221.
XX (BURN-) BURNHAM INST.
XX
PI Reed JC;
XX
DR WPI; 2001-183258/18.
XX N-PSDB; AAD02764.
PT Novel nucleic acid encoding NB-ARC and caspase associated recruitment
PT domains, used to produce polypeptides for screening for modulators of
PT apoptosis -
XX
PS Disclosure; Page 168-172; 184pp; English.
XX
CC The present sequence is a human NB-ARC and CARD containing protein (NAC)
CC beta isoform-CARD-X1 chimeric protein. NAC protein comprises a
CC nucleotide binding (NB) domain (also referred as NB-ARC domain), a
CC caspase-associated recruitment domain (CARD) and a TIM-Barrel-like
CC domain. CARD-X protein comprises a caspase-associated recruitment domain
CC (CARD) and a TIM-Barrel-like domain. The caspases, cysteine aspartyl
CC proteases, are principal effectors of apoptosis. NAC and CARD-X are used
CC for screening modulators that modulates apoptosis, cytokine production,
CC cytokine receptor signalling and other cellular processes. They can act
CC as an immunogen for the production of polyclonal and monoclonal
CC antibodies. They can also be used to diagnose and treat inflammatory

XX	Reed JC;	
XX	WPI; 2001-183258/18.	
DR	N-PSDB; AAD02765.	
XX	Novel nucleic acid encoding NB-ARC and caspase associated recruitment domains, used to produce polypeptides for screening for modulators of apoptosis -	
XX	Disclosure; Page 179-183; 184pp; English.	
XX	The present sequence is a human NB-ARC and CARD containing protein (NAC) gamma or delta isoform-CARD-X1 chimeric protein. NAC protein comprises a nucleotide binding (NB) domain (also referred as NB-ARC domain), a caspase-associated recruitment domain (CARD) and a TIM-Barrel-like domain. CARD-X protein comprises a caspase-associated recruitment domain (CARD) and a TIM-Barrel-like domain. The caspases, cysteine aspartyl proteases, are principal effectors of apoptosis. NAC and CARD-X are used for screening modulators that modulates apoptosis, cytokine production, cytokine receptor signalling and other cellular processes. They can act as an immunogen for the production of polyclonal and monoclonal antibodies. They can also be used to diagnose and treat inflammatory disorders such as sepsis, fibrosis and arthritis and cancer pathologies such as adenocarcinomas and leukaemias.	
XX	Sequence 1424 AA;	
SQ	Query Match 81.9%; Score 6174; DB 22; Length 1424;	
	Best Local Similarity 82.1%; Pred. No. 0;	
	Matches 1199; Conservative 53; Mismatches 122; Indels 86; Gaps 8;	
QY	1 MAGGAWRLACYLEFLKKELEKEFQALLANKAHSRSSGGETPAQPEKTSKMEVASYLVAQ 60	
DB	1 MAGGAWRLACYLEFLKKELEKEFQALLANKAHSRSSGGETPAQPEKTSKMEVASYLVAQ 60	
QY	61 YGQRAWDALHTWEQGLRSLCAQAEQAGHSPSPYPSPSEPHLGSPQSTSTAVLMPW 120	
DB	61 YGQRAWDALHTWEQGLRSLCAQAEQAGHSPSPYPSPSEPHLGSPQSTSTAVLMPW 120	
QY	121 IHELPAGCTGSSRRVRLQPLDTSGRWRREISALYLQALPSSPDHESQSPNAPTST 180	
DB	121 IHELPAGCTGSSRRVRLQPLDTSGRWRREISALYLQALPSSPDHESQSPNAPTST 180	
QY	181 AVLGSGSPQPSLAPREQAPGTQWPLDTSYIYTEREREREKSEKGRPPWAAVVG 240	
DB	181 AVLGSGSPQPSLAPREQAPGTQWPLDTSYIYTEREREREKSEKGRPPWAAVVG 240	
QY	241 PPOAHTSLQPHHPWPSPVRESLSTWPKNEDFNQKFTQLLLQRPSPRSDPLVKRSW 300	
DB	241 PPOAHTSLQPHHPWPSPVRESLSTWPKNEDFNQKFTQLLLQRPSPRSDPLVKRSW 300	
QY	301 PDVVEENRGLHIBIRDLFGPLDTQPRIVILQAGAGIKSTLQVKEANGRGQLYGR 360	
DB	301 PDVVEENRGLHIBIRDLFGPLDTQPRIVILQAGAGIKSTLQVKEANGRGQLYGR 360	
QY	361 FOHVYFSCRELQAKSVSLAEILGKGTATPAIRQILSRPERLLFILDGVDEPGWVLQ 420	
DB	361 FOHVYFSCRELQAKSVSLAEILGKGTATPAIRQILSRPERLLFILDGVDEPGWVLQ 420	
QY	421 EPSSELCLHWSQPPADALLGSLGKTLTILPEASFLITARTALQNLIPSLQARWVEVLG 480	
DB	421 EPSSELCLHWSQPPADALLGSLGKTLTILPEASFLITARTALQNLIPSLQARWVEVLG 480	
QY	481 FSSSRKRYFYFTDERQAIRAFRLVKSNNELWALCLVPWVSWLACTCLMQMKKEKL 540	
DB	481 FSSSRKRYFYFTDERQAIRAFRLVKSNNELWALCLVPWVSWLACTCLMQMKKEKL 540	
QY	541 TLTSKTTTTLCHYLQAQAOPLGQRLDLCSLAEGITWOKTILPSPDRLKHGLDGI 600	
DB	541 TLTSKTTTTLCHYLQAQAOPLGQRLDLCSLAEGITWOKTILPSPDRLKHGLDGI 600	
QY	601 ISTFLKMGILQEHPIPLSYFHLFCQEFFAAMSVVLEDEKGRGKHSNCIIDLKTLAY 660	

DB	601 ISTFLKMGILQEHPIPLSYFHLFCQEFFAAMSVVLEDEKGRGKHSNCIIDLKTLAY 660	
QY	661 GIHGLFGASTTRFLGLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQQLLQPHSLESILH 720	
DB	661 GIHGLFGASTTRFLGLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQQLLQPHSLESILH 720	
QY	721 CLYETRNKTFITQVNAHFEEMGCMVETDMELLVCTFCIKFSRHVKKQLQIEGRQHRSTWS 780	
DB	721 CLYETRNKTFITQVNAHFEEMGCMVETDMELLVCTFCIKFSRHVKKQLQIEGRQHRSTWS 780	
QY	781 PMVVLFWPVPVTDAYWQILFSLVKVTENLKELDLSGNSLSHSAVKSCKTLRRRCLE 840	
DB	781 PMVVLFWPVPVTDAYWQILFSLVKVTENLKELDLSGNSLSHSAVKSCKTLRRRCLE 840	
QY	841 TLRLAGCGLTAECDCKDLAFGLRANOTLTDELDSFNVLTDAGAKHLCQRLRQPSCKLRLQ 900	
DB	841 TLRLAGCGLTAECDCKDLAFGLRANOTLTDELDSFNVLTDAGAKHLCQRLRQPSCKLRLQ 900	
QY	901 LVSCGLTSDCCODLASVLSASPSLKELDLQNNLDVGVRLLCCEGLRHPACKLIRGLDQ 960	
DB	901 LVSCGLTSDCCODLASVLSASPSLKELDLQNNLDVGVRLLCCEGLRHPACKLIRGLDQ 960	
QY	961 TTLSDEMRELALAEQERPOLLI FSRRKPSVMTPEGLDTGEMSNSTSLKQRIGSERA 1020	
DB	961 TTLSDEMRELALAEQERPOLLI FSRRKPSVMTPEGLDTGEMSNSTSLKQRIGSERA 1020	
QY	1021 ASHVAQANLKLLDVSKIPPIABIAESSPEVVPVELLCVPSPASQGLHTKPLGTDDDFW 1080	
DB	991 ASHVAQANLKLLDVSKIPPIABIAESSPEVVPVELLCVPSPASQGLHTKPLGTDDDFW 1080	
QY	1081 GPTGVATEVVDKKNLYRVHPFVAGSYRWPNVTGLCFVMREAVTVVEIFCVWMDQFLG-EI 1139	
DB	1051 GPEGNVDELIDKSTNRYSVMPFTAGWYLSATGLGLVRDEVTVTIAFGSWSQHLALDL 1110	
QY	1140 NPQHSWVAGPLLDITKAEP-GAVEAVHLPHFVALQGHVDTSLFQVAHFKBEGMLLEKPA 1198	
DB	1111 QHHEQMLVGGPLFDVTAEPPEAVAIHLPFHFSIQ-GEVDVSWFLVAHPKNEGMVLEHPA 1169	
QY	1199 RVELHIVLENPSFSPGLVLLKMIHNAIRFIPVTSVLLYHRVHPVEVTFHLYLIPSDCS 1258	
DB	1170 RVEPFYAVLESFSPSLMGILLRIASGTRLSIPITNTLIYYHPHEDIKFHLIYVPSDAL 1229	
QY	1259 IR-----KELELCYRSPGEDQLF 1276	
DB	1230 LTKAIDDEEDRPHGVRLQTSPPMEPLNFGSSYIVSNSANLKVMPKELKLSYRSPGEIQHF 1289	
QY	1277 SEFYVGHLSGTRLOVKDKKDETLWEALVKPGDLMPTTLIPPARIAVPSPLDAPQLLH 1336	
DB	1290 SKFYAGQMKETQLEITEKRHGTLYMDTEVKPDLQLVAASAPP-----PFSGAA--- 1339	
QY	1337 FVDQYREQLIARVTSVEVVDKHL-GVLSQSQYERVLAEINTRPSQMRKLFSLSQSWDRK 1395	
DB	1340 FYKENHRLQARMGDLKGVLDDLQDNEVLITENEKELVBEQKTRQSKNEALLSNVEKKGDL 1399	
QY	1396 CKDGLYQALKEPHLIMEEL 1415	
DB	1400 ALDVLFRSISERDPIVLSYL 1419	

Search completed: January 29, 2004, 13:45:54
Job time : 74.5142 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2004, 03:03:41 ; Search time 1728.98 Seconds
(without alignments)
11475.528 Million cell updates/sec

Title: US-09-996-617-1
Perfect score: 5444
Sequence: 1 gccccaggcctggagaggt.....aggaataagaattacctac 5444

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 2434939 seqs, 1822278265 residues

Total number of hits satisfying chosen parameters: 4869878

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:
10: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:
11: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:
12: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:
14: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:
17: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:
18: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5444	100.0	5444	10	US-09-996-617-1
2	5444	100.0	5444	10	US-09-931-071-1
3	5444	100.0	5444	11	US-09-956-712-3
4	5284.2	97.1	6531	11	US-09-956-712-11
5	4882.4	89.7	5100	11	US-09-956-712-12
6	4287	78.7	4287	10	US-09-996-617-5
7	4287	78.7	4287	10	US-09-931-071-5
8	4148	76.2	4422	10	US-09-388-221-1
9	4100	75.3	4200	10	US-09-388-221-3
10	3958	72.7	4332	10	US-09-388-221-5
11	3400.4	62.5	4556	10	US-09-388-221-9
12	3210.4	59.0	4466	10	US-09-388-221-11
13	2572.6	47.3	2657	11	US-09-895-298-22
14	1710.8	31.4	9649	11	US-09-956-712-10
15	694.4	12.8	720	13	US-10-027-632-150437

c	16	694.4	12.8	720	14	US-10-027-632-150437	Sequence 150437,
	17	564	10.4	564	13	US-10-029-386-2806	Sequence 2806, Ap
	18	508.8	9.3	725	13	US-10-027-632-20320	Sequence 20320, A
	19	508.8	9.3	725	14	US-10-027-632-20320	Sequence 20320, A
	20	489.4	9.0	491	10	US-09-796-692-5036	Sequence 5036, Ap
	21	489.4	9.0	491	12	US-10-057-4758-5036	Sequence 5036, Ap
	22	489.4	9.0	491	12	US-10-154-884B-5036	Sequence 5036, Ap
	23	489.4	9.0	491	15	US-10-040-862-5036	Sequence 5036, Ap
c	24	484.4	8.9	578	11	US-09-764-891-1745	Sequence 1745, Ap
	25	482.8	8.9	487	15	US-10-060-036-3298	Sequence 3298, Ap
c	26	414.4	7.6	416	15	US-10-060-036-3299	Sequence 3299, Ap
	27	399.8	7.3	3186	15	US-10-066-521-17	Sequence 17, Appl
	28	398.2	7.3	3466	12	US-10-108-260A-718	Sequence 718, Appl
	29	398	7.3	3108	11	US-09-965-621-23	Sequence 23, Appl
	30	398	7.3	3108	12	US-10-407-866-23	Sequence 23, Appl
	31	398	7.3	3218	12	US-10-407-866-67	Sequence 67, Appl
	32	369.8	6.8	3102	13	US-10-132-967-6	Sequence 6, Appli
	33	369.8	6.8	3102	14	US-10-127-516-6	Sequence 6, Appli
	34	369.8	6.8	3102	14	US-10-027-629-6	Sequence 6, Appli
	35	369.8	6.8	3857	13	US-10-132-967-4	Sequence 4, Appli
	36	369.8	6.8	3857	14	US-10-127-516-4	Sequence 4, Appli
	37	369.8	6.8	3857	14	US-10-027-629-4	Sequence 4, Appli
	38	340.6	6.3	371	12	US-10-242-535A-20893	Sequence 20893, A
	39	318.4	5.8	322	10	US-09-796-692-2680	Sequence 2680, Ap
	40	318.4	5.8	322	12	US-10-057-475B-2680	Sequence 2680, Ap
	41	318.4	5.8	322	12	US-10-154-884B-2680	Sequence 2680, Ap
	42	318.4	5.8	322	15	US-10-040-862-2680	Sequence 2680, Ap
	43	313.4	5.8	2691	15	US-10-066-521-21	Sequence 21, Appl
	44	292.8	5.4	2524	11	US-09-965-621-58	Sequence 58, Appl
	45	292.8	5.4	2524	12	US-10-407-866-58	Sequence 58, Appl

ALIGNMENTS

RESULT 1

US-09-996-617-1
; Sequence 1, Application US/09996617
; Patent No. US20020128198A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; FILE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-340001
; CURRENT APPLICATION NUMBER: US/09/996,617
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: 09/931,071
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/428,252
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5444
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (523)...(4809)
US-09-996-617-1

Query Match	100.0%;	Score 5444;	DB 10;	Length 5444;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 5444;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	GCCTCAGGCGCTGGAGAGGTCTGAAGAAACCTGGAGCCAGAGCCGCGGGCTCCACTCT	60	
Db	1	GCCTCAGGCGCTGGAGAGGTCTGAAGAAACCTGGAGCCAGAGCCGCGGGCTCCACTCT	60	
Qy	61	GGTTCTGAAGCCCACTTCCCTGCTGCGGTCTCTCCACCCCACTTCTTCAGCTT	120	

Db	61	GGGTTCTGAAGACCCATTCCTGCTCTGGGCTCCTCCACCCCAACCTCTTCTCAGCGTT	120
Qy	121	GCAGCTCAAGGGTTGATCTCAGGAGTCCAGGACCCAGGAGAGGGAAGATCTGAGGAACA	180
Db	121	GCAGCTCAAGGGTTGATCTCAGGAGTCCAGGACCCAGGAGAGGGAAGATCTGAGGAACA	180
Qy	181	CAGAACAGTGAGGTTGGCCACAACCCATCTCCGTACCAATCTCCCTCACCTTCAC	240
Db	181	CAGAACAGTGAGGTTGGCCACAACCCATCTCCGTACCAATCTCCCTCACCTTCAC	240
Qy	241	CCTCCCTGCTGGCCCTGGACCCCATCCAGGACCTCCCTATCAGCTGACTTCTTCCAGT	300
Db	241	CCTCCCTGCTGGCCCTGGACCCCATCCAGGACCTCCCTATCAGCTGACTTCTTCCAGT	300
Qy	301	GTCTTTCAGGCGCTCTGGGCTCCTCCCTCGCTGGTTTTCTTACACACTCCCCCTCTAT	360
Db	301	GTCTTTCAGGCGCTCTGGGCTCCTCCCTCGCTGGTTTTCTTACACACTCCCCCTCTAT	360
Qy	361	CGGCGTCTATCTGTAGTGCCTCGGATTTATAAACTGGGTTCCGAATGCTGAATAAGA	420
Db	361	CGGCGTCTATCTGTAGTGCCTCGGATTTATAAACTGGGTTCCGAATGCTGAATAAGA	420
Qy	421	GACGTAAGACCAAGGCAAGGACGACACTGTTCTCTGCTGCTGCTATACCTCACAC	480
Db	421	GACGTAAGACCAAGGCAAGGACGACACTGTTCTCTGCTGCTGCTATACCTCACAC	480
Qy	481	CTGGCAATCCCCAGACACCTCTTAACTCCGGGACAGAGATGCGTGGCGGAGCCTGG	540
Db	481	CTGGCAATCCCCAGACACCTCTTAACTCCGGGACAGAGATGCGTGGCGGAGCCTGG	540
Qy	541	GGCGCCTGGCCTGTTACTTGGAGTTCCTGAAGAGGAGGAGCTGAAGGAGTTCACGCTT	600
Db	541	GGCGCCTGGCCTGTTACTTGGAGTTCCTGAAGAGGAGGAGCTGAAGGAGTTCACGCTT	600
Qy	601	CTGCTCGCAATAAAGCGCACTCAGGAGCTCTTCGGGTGAGACACCGCTCAGCCAGAG	660
Db	601	CTGCTCGCAATAAAGCGCACTCAGGAGCTCTTCGGGTGAGACACCGCTCAGCCAGAG	660
Qy	661	AAGACGAGTGCATGGAGTGGCGCTCGTACTGCTGGCTCAGTATGGGAGCAGCGGCC	720
Db	661	AAGACGAGTGCATGGAGTGGCGCTCGTACTGCTGGCTCAGTATGGGAGCAGCGGCC	720
Qy	721	TGGACCTTAGCCCTCCATACCTGGGAGCAGATGGGGCTGAGGTCACTGTGCGCCAAAGCC	780
Db	721	TGGACCTTAGCCCTCCATACCTGGGAGCAGATGGGGCTGAGGTCACTGTGCGCCAAAGCC	780
Qy	781	CAGGAAGGGGAGGCGACTCTCCCTCATTTCCCTTACAGCCCAAGTGAACCCCACTGGG	840
Db	781	CAGGAAGGGGAGGCGACTCTCCCTCATTTCCCTTACAGCCCAAGTGAACCCCACTGGG	840
Qy	841	TCTCCAGCCAAACCACTCCACCGCAGTGTAAATGCCCTGGATCCATGAATTCGCGCG	900
Db	841	TCTCCAGCCAAACCACTCCACCGCAGTGTAAATGCCCTGGATCCATGAATTCGCGCG	900
Qy	901	GGGTGCACCCAGGGCTCAGAGAGAAGGGTTTTGAGACAGCTGCTGCATCATCTGACGC	960
Db	901	GGGTGCACCCAGGGCTCAGAGAGAAGGGTTTTGAGACAGCTGCTGCATCATCTGACGC	960
Qy	961	CGCTGGAGAGAAATCTCTGCTCTACTCTCTTACCAAGCTTTTCAAGCTTCCCAAGACAT	1020
Db	961	CGCTGGAGAGAAATCTCTGCTCTACTCTCTTACCAAGCTTTTCAAGCTTCCCAAGACAT	1020
Qy	1021	GAGTCTCCAGCCAGGAGTCAACCAACCCGCCCACTATCCACAGCTGCTGGGAGCTGG	1080
Db	1021	GAGTCTCCAGCCAGGAGTCAACCAACCCGCCCACTATCCACAGCTGCTGGGAGCTGG	1080
Qy	1081	GGATCCCCACCTCAGCCAGCCTTAGCACCCAGAGAGCAGGAGTCTCTGGGACCCAATGG	1140
Db	1081	GGATCCCCACCTCAGCCAGCCTTAGCACCCAGAGAGCAGGAGTCTCTGGGACCCAATGG	1140
Qy	1141	CCTCTGGATGAAACGTCTAGGAATTTTACTTACACAGAAATCAGAGAAAGAGAGAGAAA	1200
Db	1141	CCTCTGGATGAAACGTCTAGGAATTTTACTTACACAGAAATCAGAGAAAGAGAGAGAAA	1200

[illegible]

QY 2281 AGTCAGATGACCTCAGGAAGCATGGTTAGATGGGGCCATCATCTCCACCTCTTGAAG 2340
 Db 2281 AGTCAGATGACCTCAGGAAGCATGGTTAGATGGGGCCATCATCTCCACCTCTTGAAG 2340
 QY 2341 ATGGGTATTCTTCAAGACAGACCCCATCCCTCTGAGCTACAGCTTCACTACCTCTGTTTC 2400
 Db 2341 ATGGGTATTCTTCAAGACAGACCCCATCCCTCTGAGCTACAGCTTCACTACCTCTGTTTC 2400
 QY 2401 CAAGAGTTCTTTGCAGCAATGTCCTATGTTCTGGAGGATGAGAAGGGGAGAGGTAACAT 2460
 Db 2401 CAAGAGTTCTTTGCAGCAATGTCCTATGTTCTGGAGGATGAGAAGGGGAGAGGTAACAT 2460
 QY 2461 TCTAATTGATCATAGATTGGAAAAGACGCTAGAAAGCATATGAAATACATGCGCTGTTT 2520
 Db 2461 TCTAATTGATCATAGATTGGAAAAGACGCTAGAAAGCATATGAAATACATGCGCTGTTT 2520
 QY 2521 GGGGATCAACACAGTTTCTATTTGGGCCCTGTTAAGTGATGAGGGGAGAGAGATG 2580
 Db 2521 GGGGATCAACACAGTTTCTATTTGGGCCCTGTTAAGTGATGAGGGGAGAGAGATG 2580
 QY 2581 GAGAACATCTTTCACTGCCGGTGTCTCAGGGGAGGAACTGATGCACTGGGTCCGCTCC 2640
 Db 2581 GAGAACATCTTTCACTGCCGGTGTCTCAGGGGAGGAACTGATGCACTGGGTCCGCTCC 2640
 QY 2641 CTGCACTGCTGTCAGCCACCACTCTCTGGAGTCCCTCCACTGCTTGTGACAGACTCGG 2700
 Db 2641 CTGCACTGCTGTCAGCCACCACTCTCTGGAGTCCCTCCACTGCTTGTGACAGACTCGG 2700
 QY 2701 ACACAAAGCTTCTGACACAGTATGGGCCATTTTGAAGAAATGGGCATGTTGTAGAA 2760
 Db 2701 ACACAAAGCTTCTGACACAGTATGGGCCATTTTGAAGAAATGGGCATGTTGTAGAA 2760
 QY 2761 ACAGACATGGAGCTCTTAGTGTGCATTTCTGCAATTAATTAATTCAGCCGCCACCTGAAGAG 2820
 Db 2761 ACAGACATGGAGCTCTTAGTGTGCATTTCTGCAATTAATTAATTCAGCCGCCACCTGAAGAG 2820
 QY 2821 CTTCACTGATTTGAGGGCAGGACGACAGATCAACATGGAGCCCAACCATGTAFTCTG 2880
 Db 2821 CTTCACTGATTTGAGGGCAGGACGACAGATCAACATGGAGCCCAACCATGTAFTCTG 2880
 QY 2881 TTCAAGTGGTCCAGTACAGATGCTATTGGCAGATTTCTTCTCCGTCTCAAGTTC 2940
 Db 2881 TTCAAGTGGTCCAGTACAGATGCTATTGGCAGATTTCTTCTCCGTCTCAAGTTC 2940
 QY 2941 ACCAGAAACCTGAAGAGCTGACCTAAGTGAACCTCGCTGAGCCACTCTGCAGTGAAG 3000
 Db 2941 ACCAGAAACCTGAAGAGCTGACCTAAGTGAACCTCGCTGAGCCACTCTGCAGTGAAG 3000
 QY 3001 AGTCTTTGTAAGACCCCTGAGACGCGCTCGCTGCTGAGACCCCTGCGGTTGGCTGCG 3060
 Db 3001 AGTCTTTGTAAGACCCCTGAGACGCGCTCGCTGCTGAGACCCCTGCGGTTGGCTGCG 3060
 QY 3061 TGTGGCTCAGAGCTGAGACTGCAAGGACCTTGCCTTTGGGCTGAGACCCAAACGAC 3120
 Db 3061 TGTGGCTCAGAGCTGAGACTGCAAGGACCTTGCCTTTGGGCTGAGACCCAAACGAC 3120
 QY 3121 CTGACCGAGCTGAGCTGAGCTCAATGCTCAGGATGCTGAGCCAAACACCTTTGC 3180
 Db 3121 CTGACCGAGCTGAGCTGAGCTCAATGCTCAGGATGCTGAGCCAAACACCTTTGC 3180
 QY 3181 CAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCAGCTGAGCTGCTGAGCTGTCGCTC 3240
 Db 3181 CAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCAGCTGAGCTGCTGAGCTGTCGCTC 3240
 QY 3241 AGTCTGACTGTCGAGACCTGCGCTCTGCTGCTTGTAGTCCAGCCCAAGCCTGAAGAG 3300
 Db 3241 AGTCTGACTGTCGAGACCTGCGCTCTGCTGCTTGTAGTCCAGCCCAAGCCTGAAGAG 3300
 QY 3301 CTAGACCTCAGCAGAACCACTGGAATGAGTTGGCTGCGACTGCTCTGTCAGGGGCTC 3360
 Db 3301 CTAGACCTCAGCAGAACCACTGGAATGAGTTGGCTGCGACTGCTCTGTCAGGGGCTC 3360
 QY 3361 AGGCATCTCGCTGCAAACTCATACGCTGGGCTGGACAGACAACTCTGATGATGAG 3420

Db 3361 AGGCATCTCGCTGCAAACTCATACGCTGGGCTGGACAGACAACTCTGATGATGAG 3420
 QY 3421 ATGAGGACAGAACTGAGGGCCCTGGAGCAGGAGAAACCTCAGCTGCTCATCTTCAAGAGA 3480
 Db 3421 ATGAGGACAGAACTGAGGGCCCTGGAGCAGGAGAAACCTCAGCTGCTCATCTTCAAGAGA 3480
 QY 3481 CGSAAACCAAGTGTGATGACCCCTA CTGAGGCGCTGGATACGGGAGAGATGAGTAATAGC 3540
 Db 3481 CGSAAACCAAGTGTGATGACCCCTA CTGAGGCGCTGGATACGGGAGAGATGAGTAATAGC 3540
 QY 3541 ACATCTCACTCAAGCGGACAGACTCGGATCAGAGGGCGGCTTCCATGTTGCTCAG 3600
 Db 3541 ACATCTCACTCAAGCGGACAGACTCGGATCAGAGGGCGGCTTCCATGTTGCTCAG 3600
 QY 3601 GCTAATCTCAAACTCTGACGCTGAGCAAGATCTTCCCAATTTGCTGAGATTCGACAGGAA 3660
 Db 3601 GCTAATCTCAAACTCTGACGCTGAGCAAGATCTTCCCAATTTGCTGAGATTCGACAGGAA 3660
 QY 3661 AGCTCCCGCAGAGTAGTACCGGTGGAACTCTTTGTCGTCCTTCTCTGCTCTCAAAGG 3720
 Db 3661 AGCTCCCGCAGAGTAGTACCGGTGGAACTCTTTGTCGTCCTTCTCTGCTCTCAAAGG 3720
 QY 3721 GACCTGCATACGAAGCTTTTGGGACCTGACATGACTCTTGGGGCCCCACGGGGCTGTG 3780
 Db 3721 GACCTGCATACGAAGCTTTTGGGACCTGACATGACTCTTGGGGCCCCACGGGGCTGTG 3780
 QY 3781 GCTACTGAGGTAGTTGACAAAGAAAGAACTTTGACCGAGTTCACTTCCCTGTAGCTGCG 3840
 Db 3781 GCTACTGAGGTAGTTGACAAAGAAAGAACTTTGACCGAGTTCACTTCCCTGTAGCTGCG 3840
 QY 3841 TCCTACCGCTGGCCCAACA CGGGTCTCTGCTTTGTGATGAGAGAGCGGTGACCGTTGAG 3900
 Db 3841 TCCTACCGCTGGCCCAACA CGGGTCTCTGCTTTGTGATGAGAGAGCGGTGACCGTTGAG 3900
 QY 3901 ATTGAATTTCTGTGTGGGACAGTTCTCTGGGTGAGATCAACCCACAGCAGCTGGATG 3960
 Db 3901 ATTGAATTTCTGTGTGGGACAGTTCTCTGGGTGAGATCAACCCACAGCAGCTGGATG 3960
 QY 3961 GTGGCAGGCGCTCTGTCTGACATCAAGGCTGAGCTGAGCTGTGGAAGCTGTGACACCTC 4020
 Db 3961 GTGGCAGGCGCTCTGTCTGACATCAAGGCTGAGCTGAGCTGTGGAAGCTGTGACACCTC 4020
 QY 4021 CCTCACTTTGTGGTCTCTCAAGGGGGCCATGTGGAACA CATCCCTGTTCCAAATGSCCCAC 4080
 Db 4021 CCTCACTTTGTGGTCTCTCAAGGGGGCCATGTGGAACA CATCCCTGTTCCAAATGSCCCAC 4080
 QY 4081 TTTAAAGAGGAGGGATGCTCTCTGAGAAAGCCAGCCAGGGTGGAGCTGCATCATAGTT 4140
 Db 4081 TTTAAAGAGGAGGGATGCTCTCTGAGAAAGCCAGCCAGGGTGGAGCTGCATCATAGTT 4140
 QY 4141 CTGGAAAACCCAGCTTCTCCCTCTGGAGTCTCTGGAATAATGATCCATAATGCCCTG 4200
 Db 4141 CTGGAAAACCCAGCTTCTCCCTCTGGAGTCTCTGGAATAATGATCCATAATGCCCTG 4200
 QY 4201 CGCTTCACTCCCGTCACCTCTGTGTGTGTTTACACCGCGTCCATCTGAGGAAAGTC 4260
 Db 4201 CGCTTCACTCCCGTCACCTCTGTGTGTGTTTACACCGCGTCCATCTGAGGAAAGTC 4260
 QY 4261 ACCTTCCAACCTCTA CCGTATCCCAAGTGAAGTCTCCATTCGGAAGGAACTGGAGCTGCG 4320
 Db 4261 ACCTTCCAACCTCTA CCGTATCCCAAGTGAAGTCTCCATTCGGAAGGAACTGGAGCTGCG 4320
 QY 4321 TATCGAAGCCCTGGAGAGACAGCTGTTCTCGGAGTCTTACGTTGGCAGCTTGGATCA 4380
 Db 4321 TATCGAAGCCCTGGAGAGACAGCTGTTCTCGGAGTCTTACGTTGGCAGCTTGGATCA 4380
 QY 4381 GGGATCAGGCTGCAAGTGAAGAAAGATGAGACTCTGCTGTGGAGGCTTGGTG 4440
 Db 4381 GGGATCAGGCTGCAAGTGAAGAAAGATGAGACTCTGCTGTGGAGGCTTGGTG 4440
 QY 4441 AAACAGGAGATCTCATGCTGCAACTACTCTGATCCCTCCAGCCCGATAGCCGTA CTT 4500

4441 AAACAGGAGATCTCATGCTGCAACTACTCTGATCCCTCCAGCCGCATAGCCGTACCT 4500
4501 TCACCTCTGGATGCCCGGAGTGTCTGCACTTTGTGACACAGTATCGAGAGAGCTGATA 4560
4501 TCACCTCTGGATGCCCGGAGTGTCTGCACTTTGTGACACAGTATCGAGAGAGCTGATA 4560
4561 GCCCGAGTGACATCGGTGGAGTGTCTTGGACAAACTCGATGGACAGGTGTCTGAGCCAG 4620
4561 GCCCGAGTGACATCGGTGGAGTGTCTTGGACAAACTCGATGGACAGGTGTCTGAGCCAG 4620
4621 GAGCAGTACAGAGAGTGTCTGAGTGTGAGAACACAGAGGCCAGCAGATGCGGAAGCTGTTT 4680
4621 GAGCAGTACAGAGAGTGTCTGAGTGTGAGAACACAGAGGCCAGCAGATGCGGAAGCTGTTT 4680
4681 AGCTTGGAGTGTCTGAGTGTGAGAACACAGAGGCCAGCAGATGCGGAAGCTGTTT 4740
4681 AGCTTGGAGTGTCTGAGTGTGAGAACACAGAGGCCAGCAGATGCGGAAGCTGTTT 4740
4741 ACCATCTCTCACCTCATTTATGGAATCTCTGGAGAGGCGAGCAAAAGGAGCTCTTGCCA 4800
4741 ACCATCTCTCACCTCATTTATGGAATCTCTGGAGAGGCGAGCAAAAGGAGCTCTTGCCA 4800
4801 CTCAGAGTGAAGTATCAACACAGCCCTTGACCCCTTGAGTCTGCTGCTGCTGAGCC 4860
4801 CTCAGAGTGAAGTATCAACACAGCCCTTGACCCCTTGAGTCTGCTGCTGCTGAGCC 4860
4861 TTCTTTGGGTCTCAGTCTTCTTCTGCAAAAGTGTGCAATCTGGTTTGCCTTCCAGCA 4920
4861 TTCTTTGGGTCTCAGTCTTCTTCTGCAAAAGTGTGCAATCTGGTTTGCCTTCCAGCA 4920
4921 CTAAAGTATGGAATCTTCTGATGATGCTTCTGCTGGGCAATATGTGCTCCATGCGAGGATG 4980
4921 CTAAAGTATGGAATCTTCTGATGATGCTTCTGCTGGGCAATATGTGCTCCATGCGAGGATG 4980
4981 CCACAGGGGGCCCCAGTCCAGGTGGCTTAAACAGCATCTCAGGGAATGTCCATCTGGAGCT 5040
4981 CCACAGGGGGCCCCAGTCCAGGTGGCTTAAACAGCATCTCAGGGAATGTCCATCTGGAGCT 5040
5041 GGCAGACCCCTGCGAGACCTCATAGACCTCATCTGCTGGTGGCCACAGCAGCCAGCTAGA 5100
5041 GGCAGACCCCTGCGAGACCTCATAGACCTCATCTGCTGGTGGCCACAGCAGCCAGCTAGA 5100
5101 GGCCTCCGATCCCATCCAGGGGCAAGAGGATAGGAGGACATGGAAACCAATTTGCCCTC 5160
5101 GGCCTCCGATCCCATCCAGGGGCAAGAGGATAGGAGGACATGGAAACCAATTTGCCCTC 5160
5161 TGCTGTGTCAAGGTGAGCCCAAAATTTGGGTTTCAAGGTGGAGGCCACGTGGATTC 5220
5161 TGCTGTGTCAAGGTGAGCCCAAAATTTGGGTTTCAAGGTGGAGGCCACGTGGATTC 5220
5221 TTGGCTTTGTAAGGAGATCTACAGAGCAGCCCAAGAGTAAAGTGGAGGATTT 5280
5221 TTGGCTTTGTAAGGAGATCTACAGAGCAGCCCAAGAGTAAAGTGGAGGATTT 5280
5281 ATTACAGAAATAAGGAGTATCACAGCTCTTTTAGAATTTGTCTAGCAGCTTTCCAGTT 5340
5281 ATTACAGAAATAAGGAGTATCACAGCTCTTTTAGAATTTGTCTAGCAGCTTTCCAGTT 5340
5341 TTACAGAAAAACCCCTATAAATTTAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 5400
5341 TTACAGAAAAACCCCTATAAATTTAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 5400
5401 ACMAAAGAAAAATGAAATTAAGGAATTAAGGATTAACCTAC 5444
5401 ACMAAAGAAAAATGAAATTAAGGAATTAAGGATTAACCTAC 5444

RESULT 2
US-09-931-071-1
; Sequence 1, Application US/09931071
; Patent No. US20020128219A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John

; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; FILE REFERENCE: 07334-335001
; CURRENT APPLICATION NUMBER: US/09/931,071
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/428,252
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5444
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (523)...(4809)
US-09-931-071-1

Query Match 100.0%; Score 5444; DB 10; Length 5444;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCAGGGCCTGGAGAGGTCTGAAGAAACCTGGAGCCAGCAGCCCGGGGCTCCACTCT 60
Db 1 GCCCAGGGCCTGGAGAGGTCTGAAGAAACCTGGAGCCAGCAGCCCGGGGCTCCACTCT 60
QY 61 GGGTCTGAAGACCCATTTCTGCTGGGTCTCTCCACCCCACTTTCTGAGCCTT 120
Db 61 GGGTCTGAAGACCCATTTCTGCTGGGTCTCTCCACCCCACTTTCTGAGCCTT 120
QY 121 GCAGCTCAAGGTTGATCTCAGGAGTCCAGACCCAGAGAGGAGGAAGTCTGAGGAACA 180
Db 121 GCAGCTCAAGGTTGATCTCAGGAGTCCAGACCCAGAGAGGAGGAAGTCTGAGGAACA 180
QY 181 CAGAACAGTGAAGCTTGGCCACACCCCACTCTCCCTCACCACATCTCCCTCACCCTCAC 240
Db 181 CAGAACAGTGAAGCTTGGCCACACCCCACTCTCCCTCACCACATCTCCCTCACCCTCAC 240
QY 241 CTTCTCCCTGGCCCTGGACCCCATCCAGACCTCCCTATCAGTCACTTCTTCCAGT 300
Db 241 CTTCTCCCTGGCCCTGGACCCCATCCAGACCTCCCTATCAGTCACTTCTTCCAGT 300
QY 301 GTCTTGGAGGCCCTCTGGGTCTCTCCCTCCCTGGCTTTTCCCTACCACTCCCTCTAT 360
Db 301 GTCTTGGAGGCCCTCTGGGTCTCTCCCTCCCTGGCTTTTCCCTACCACTCCCTCTAT 360
QY 361 CGGCTCTATCTGTAGTGGCTTGGGATTTATAAACTGGGTTCCGAATGCTGAATAAGA 420
Db 361 CGGCTCTATCTGTAGTGGCTTGGGATTTATAAACTGGGTTCCGAATGCTGAATAAGA 420
QY 421 GACGTTAAGAGCAAGGCAAGGACAGCAGTCTTCTGCTGCTGCTGATACCTCACCAC 480
Db 421 GACGTTAAGAGCAAGGCAAGGACAGCAGTCTTCTGCTGCTGCTGATACCTCACCAC 480
QY 481 CTGGGAACATCCCCAGACACCCCTTTAACTCCGGAAGAGAGTGGCTGGGAGGAGCTGG 540
Db 481 CTGGGAACATCCCCAGACACCCCTTTAACTCCGGAAGAGAGTGGCTGGGAGGAGCTGG 540
QY 541 GGGCCCTCTGTTTACTTGGAGTCTCTGAAGAGGAGGAGCTGAAGAGTTCAGCTT 600
Db 541 GGGCCCTCTGTTTACTTGGAGTCTCTGAAGAGGAGGAGCTGAAGAGTTCAGCTT 600
QY 601 CTGCTCGCAATAAAGCCACTCCAGAGAGTCTTCCGGGTGAGACACCCGCTCAGCCAGAG 660
Db 601 CTGCTCGCAATAAAGCCACTCCAGAGAGTCTTCCGGGTGAGACACCCGCTCAGCCAGAG 660
QY 661 AAGACGAGTGGAGTGGGCTCGTACCTGCTGGCTCAGTATGGGAGGAGCGGCGCC 720
Db 661 AAGACGAGTGGAGTGGGCTCGTACCTGCTGGCTCAGTATGGGAGGAGCGGCGCC 720
QY 721 TGGGACCTAGCCCTCCATACCTGGAGCAGATGGGGCTGAGGTCACTGTGGCCCAAGCC 780
Db 721 TGGGACCTAGCCCTCCATACCTGGAGCAGATGGGGCTGAGGTCACTGTGGCCCAAGCC 780

Db 721 TGGGACCTTAGCCCTCCATACCTGGGAGCAGATGGGGCTGAGGTCACTGTGCGCCCAAGCC 780
QY 781 CAGGAAGGGGCAAGCCCACTCTCCCTCATTTCCCTACAGCCCAAGTGAACCCCACTGGGG 840
Db 781 CAGGAAGGGGCAAGCCCACTCTCCCTCATTTCCCTACAGCCCAAGTGAACCCCACTGGGG 840
QY 841 TCTCCAGCAACCCCACTCCAGCGCAGTGTCTAATGSCCTGTGATCCATGAATGCGCGG 900
Db 841 TCTCCAGCAACCCCACTCCAGCGCAGTGTCTAATGSCCTGTGATCCATGAATGCGCGG 900
QY 901 GGGTGACCCAGGGCTCAGAGAGAGGGTTTGGAGACGCTGCGTGACACATCTGAGCGC 960
Db 901 GGGTGACCCAGGGCTCAGAGAGAGGGTTTGGAGACGCTGCGTGACACATCTGAGCGC 960
QY 961 CGCTGGAGAGAAATCTCTCGCTCACTCTCTACCAAGCTCTTCCAGCTTCCCAAGCACT 1020
Db 961 CGCTGGAGAGAAATCTCTCGCTCACTCTCTACCAAGCTCTTCCAGCTTCCCAAGCACT 1020
QY 1021 GAGTCTCCAGCGGAGGTCA CCGAAGCGCCCGCACATCCACAGCAGTGTCTGGGAGCTGG 1080
Db 1021 GAGTCTCCAGCGGAGGTCA CCGAAGCGCCCGCACATCCACAGCAGTGTCTGGGAGCTGG 1080
QY 1081 GGTATCCCACTCAGCCCGCTAGCACCCAGAGCAGGAGCTCTGGGACCCCAATGG 1140
Db 1081 GGTATCCCACTCAGCCCGCTAGCACCCAGAGCAGGAGCTCTGGGACCCCAATGG 1140
QY 1141 CCTCTGGATGAACCGTCAAGAAATTTACTACACAGAAATCAGAGAAAGAGAGAGAGAA 1200
Db 1141 CCTCTGGATGAACCGTCAAGAAATTTACTACACAGAAATCAGAGAAAGAGAGAGAGAA 1200
QY 1201 TCAGAGAAAGGAGGCCCCATGGGAGCGGTGGTAGAACGCCCCCA CAGGCGCACCC 1260
Db 1201 TCAGAGAAAGGAGGCCCCATGGGAGCGGTGGTAGAACGCCCCCA CAGGCGCACACC 1260
QY 1261 AGCTACAGCCCA CACACACCGCTGTCTGTGAGAGAGAGCTCTGTGTCCACA 1320
Db 1261 AGCTACAGCCCA CACACACCGCTGTCTGTGAGAGAGAGCTCTGTGTCCACA 1320
QY 1321 TGGCCCTGGAATAATGAGGATTTTAAACAAATAATCA CAGCTGTCTACTTACAAAGA 1380
Db 1321 TGGCCCTGGAATAATGAGGATTTTAAACAAATAATCA CAGCTGTCTACTTACAAAGA 1380
QY 1381 CCTCACCCAGAGCAACCAATCCCTGGTCAAGAGAGCTGGCCCTGATTTATGTGGAGGAG 1440
Db 1381 CCTCACCCAGAGCAACCAATCCCTGGTCAAGAGAGCTGGCCCTGATTTATGTGGAGGAG 1440
QY 1441 AATCGAGGACATTTAATTTGAGATCAGAGACTTTATTTGGCCAGGCTGGATACCCAGAA 1500
Db 1441 AATCGAGGACATTTAATTTGAGATCAGAGACTTTATTTGGCCAGGCTGGATACCCAGAA 1500
QY 1501 CCTCGCATAGTCATCTG CAGGGGGCTGTGGAAATTTGGGAAGTCAACACTGGCCAGGCAG 1560
Db 1501 CCTCGCATAGTCATCTG CAGGGGGCTGTGGAAATTTGGGAAGTCAACACTGGCCAGGCAG 1560
QY 1561 GTGAAGGAAGCCCTGGGGGAGGCCAGCTGTGTATGGGGACCGCTTC CAGCATGTCTTCTAC 1620
Db 1561 GTGAAGGAAGCCCTGGGGGAGGCCAGCTGTGTATGGGGACCGCTTC CAGCATGTCTTCTAC 1620
QY 1621 TTCAGCTGCAGAGAGCTGGCCCAAGTCCAGGTGGTGAGTCTCGCTGAGCTCATCGGAAA 1680
Db 1621 TTCAGCTGCAGAGAGCTGGCCCAAGTCCAGGTGGTGAGTCTCGCTGAGCTCATCGGAAA 1680
QY 1681 GATGGGACAGCCACTCGGGCTCCATTTAGACATCTCTGTAGGCCAGCGGCTGCTC 1740
Db 1681 GATGGGACAGCCACTCGGGCTCCATTTAGACATCTCTGTAGGCCAGCGGCTGCTC 1740
QY 1741 TTTATCTCTGATGGTGTAGATGAGCCAGATGGGTCTTTCAGGAGCCGAGTTCTGAGCTC 1800
Db 1741 TTTATCTCTGATGGTGTAGATGAGCCAGATGGGTCTTTCAGGAGCCGAGTTCTGAGCTC 1800
QY 1801 TGTCTGACTTGGAGCCAGCCAGCGGGGATGCACTGCTGGGCAAGTTTGTCTGGGAAA 1860
Db 1801 TGTCTGACTTGGAGCCAGCCAGCGGGGATGCACTGCTGGGCAAGTTTGTCTGGGAAA 1860

QY 1861 ACTATACTTCCGAGGCACTCTCTGATCA CGGCTCGGACCA CAGCTCTG CAGAACCTC 1920
Db 1861 ACTATACTTCCGAGGCACTCTCTGATCA CGGCTCGGACCA CAGCTCTG CAGAACCTC 1920
QY 1921 ATTCCTCTTTTGGAGCAGCGACGTTGGGTAGAGTCTCTGGGGTCTCTGAGTCCAGCAGG 1980
Db 1921 ATTCCTCTTTTGGAGCAGCGACGTTGGGTAGAGTCTCTGGGGTCTCTGAGTCCAGCAGG 1980
QY 1981 AAGGAATATTTCTACAGATATTTTCA CAGATGAAGGCAAGCAATTTAGAGCCTTTAGGTG 2040
Db 1981 AAGGAATATTTCTACAGATATTTTCA CAGATGAAGGCAAGCAATTTAGAGCCTTTAGGTG 2040
QY 2041 GTCAAAATCAAA CAAAGAGCTCTGGCCCTGTGTCTTGTGCCCCCTGGGTGTCTGGGTGCCC 2100
Db 2041 GTCAAAATCAAA CAAAGAGCTCTGGCCCTGTGTCTTGTGCCCCCTGGGTGTCTGGGTGCCC 2100
QY 2101 TGCACCTTGCCTGATGCAGCAGATGAAGCGGAAGGAAAACTCACACTGACTTCCAGACC 2160
Db 2101 TGCACCTTGCCTGATGCAGCAGATGAAGCGGAAGGAAAACTCACACTGACTTCCAGACC 2160
QY 2161 ACCACAA CCGCTCTGTCTACATTTACCTTGC CCAAGCTCTCCAAGCTCAGCATTGGGACC 2220
Db 2161 ACCACAA CCGCTCTGTCTACATTTACCTTGC CCAAGCTCTCCAAGCTCAGCATTGGGACC 2220
QY 2221 CAGCTCAGAGACCTCTCTCTCTCTGCTCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTTTC 2280
Db 2221 CAGCTCAGAGACCTCTCTCTCTCTGCTCTCTGCTCTGCTGCTGCTGCTGCTGCTTTC 2280
QY 2281 AGTCAGATGACCTCAGAAAGCATGGGTTAGATGGGGCCATCATCTCCACCTTCTTGAAG 2340
Db 2281 AGTCAGATGACCTCAGAAAGCATGGGTTAGATGGGGCCATCATCTCCACCTTCTTGAAG 2340
QY 2341 ATGGGTATTTCTTCAAGAGCACCCCATCCCTCTCTGAGCTACAGCTTCACTCACCTCTGTTTC 2400
Db 2341 ATGGGTATTTCTTCAAGAGCACCCCATCCCTCTCTGAGCTACAGCTTCACTCACCTCTGTTTC 2400
QY 2401 CAAAGAGTTCTTTTGCAGCAATGTCTATGCTTCTTGGAGGATGAGAA GGGAGAGGTAAACAT 2460
Db 2401 CAAAGAGTTCTTTTGCAGCAATGTCTATGCTTCTTGGAGGATGAGAA GGGAGAGGTAAACAT 2460
QY 2461 TCTAATTTGCATCATAGATTTTGGAAAGACGCTAGAAAGCATATGGAATACATGCGCTGTTT 2520
Db 2461 TCTAATTTGCATCATAGATTTTGGAAAGACGCTAGAAAGCATATGGAATACATGCGCTGTTT 2520
QY 2521 GGGGCATCAAC CACACGTTTCTTATTTGGCCCTGTTAAGTGATGAGGGGAGAGAGATG 2580
Db 2521 GGGGCATCAAC CACACGTTTCTTATTTGGCCCTGTTAAGTGATGAGGGGAGAGAGATG 2580
QY 2581 GAGAACATCTTTTCACTGCGGCTGTCTCAGGGGAGGAACTCTGATG CAGTGGGTCCCGTCC 2640
Db 2581 GAGAACATCTTTTCACTGCGGCTGTCTCAGGGGAGGAACTCTGATG CAGTGGGTCCCGTCC 2640
QY 2641 CTG CAGCTGCTGCTGCAGCCACACTCTCTGAGTCCCTCCACTGCTGTGTACAGACTCGG 2700
Db 2641 CTG CAGCTGCTGCTGCAGCCACACTCTCTGAGTCCCTCCACTGCTGTGTACAGACTCGG 2700
QY 2701 AACAAA CTTTCTGCACACAGTGTGCCCCCAATTTGAGAGAAATGGGCATGTGTAGAA 2760
Db 2701 AACAAA CTTTCTGCACACAGTGTGCCCCCAATTTGAGAGAAATGGGCATGTGTAGAA 2760
QY 2761 ACAGACATGGAGCTCTTAGTGTGCACTTTCTGCAATTTAAATTCAGCGCCCACTGGAAGAG 2820
Db 2761 ACAGACATGGAGCTCTTAGTGTGCACTTTCTGCAATTTAAATTCAGCGCCCACTGGAAGAG 2820
QY 2821 CTTT CAGCTGATTGAGGGCAGCAGACAGATCAACATGAGGCCCA CCAATGTTAGTCTCTG 2880
Db 2821 CTTT CAGCTGATTGAGGGCAGCAGCAGATCAACATGAGGCCCA CCAATGTTAGTCTCTG 2880
QY 2881 TTCAGGTGGGTCCCAAGTGCAGATGCTATTGGGAGATTTCTCTTCTCCGTCTCTCAAGGTC 2940
Db 2881 TTCAGGTGGGTCCCAAGTGCAGATGCTATTGGGAGATTTCTCTTCTCCGTCTCTCAAGGTC 2940

QY 2941 ACCAGAAACCTGAAGAGCTGGACCTAAGTGGAAACTCGCTGAGCCAATCTCGAGTGAAG 3000
 Db 2941 ACCAGAAACCTGAAGAGCTGGACCTAAGTGGAAACTCGCTGAGCCAATCTCGAGTGAAG 3000
 QY 3001 AGTCTTTTGAAGACCTTGAGACGCCCTCGCTGCTCTCTGAGACCCCTGCGGTGGCTGCG 3060
 Db 3001 AGTCTTTTGAAGACCTTGAGACGCCCTCGCTGCTCTCTGAGACCCCTGCGGTGGCTGCG 3060
 QY 3061 TGTGGCTCTCACAGCTGAGAGCTGCAAGGACCTTGGCTTGGGCTGAGAGCCAAACAGACCC 3120
 Db 3061 TGTGGCTCTCACAGCTGAGAGCTGCAAGGACCTTGGCTTGGGCTGAGAGCCAAACAGACCC 3120
 QY 3121 CTGACCGAGCTGAGACCTGAGCTTCAATGTGCTACCGGATGCTGGAGCCAAACACCTTTGC 3180
 Db 3121 CTGACCGAGCTGAGACCTGAGCTTCAATGTGCTACCGGATGCTGGAGCCAAACACCTTTGC 3180
 QY 3181 CAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCGACTGAGCTGCTGAGCTGCTGAGCTGCTG 3240
 Db 3181 CAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCGACTGAGCTGCTGAGCTGCTGAGCTGCTG 3240
 QY 3241 AGTCTGACTGCTGCGCAGACCTTGGCTCTGCTGCTTAGTGGCCAGCCCTGAGCGCTGAGGAG 3300
 Db 3241 AGTCTGACTGCTGCGCAGACCTTGGCTCTGCTGCTTAGTGGCCAGCCCTGAGCGCTGAGGAG 3300
 QY 3301 CTAGACCTGCGCAGACCAACCTTGGATGACGTTGGCTGCGACTGCTCTGAGGGGCTC 3360
 Db 3301 CTAGACCTGCGCAGACCAACCTTGGATGACGTTGGCTGCGACTGCTCTGAGGGGCTC 3360
 QY 3361 AGGCACTCTGCGTGAACCTATACGCTGCGGCTGCGACGCTGCTGAGGGGCTC 3420
 Db 3361 AGGCACTCTGCGTGAACCTATACGCTGCGGCTGCGACGCTGCTGAGGGGCTC 3420
 QY 3421 ATGAGCGAGAACTGAGGCGCTTGGAGCAGGAGAACTCAGCTGCTATCTTACGACA 3480
 Db 3421 ATGAGCGAGAACTGAGGCGCTTGGAGCAGGAGAACTCAGCTGCTATCTTACGACA 3480
 QY 3481 CGGAAACCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3540
 Db 3481 CGGAAACCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3540
 QY 3541 ACATCTCTCACTAAGCGGAGAGACTCGGATCAGAGAGGCGGCTTCCCATGTTGCTCAG 3600
 Db 3541 ACATCTCTCACTAAGCGGAGAGACTCGGATCAGAGAGGCGGCTTCCCATGTTGCTCAG 3600
 QY 3601 GCTAATCTCAAACTCTGAGACGCTGAGACAGTCTTCCCAATGCTCAGATTCAGAGGAA 3660
 Db 3601 GCTAATCTCAAACTCTGAGACGCTGAGACAGTCTTCCCAATGCTCAGATTCAGAGGAA 3660
 QY 3661 AGCTCCCGCAGAGTGTAGTACCGGTGGAATCTTGTGCGTGCCTTCTCCTGCTCTCAAGGG 3720
 Db 3661 AGCTCCCGCAGAGTGTAGTACCGGTGGAATCTTGTGCGTGCCTTCTCCTGCTCTCAAGGG 3720
 QY 3721 GACCTGCATACGAAGCCTTTGGGACTGACGATGCTTGGGGCCCCACGGGCGCTGTG 3780
 Db 3721 GACCTGCATACGAAGCCTTTGGGACTGACGATGCTTGGGGCCCCACGGGCGCTGTG 3780
 QY 3781 GCTACTGAGTGTGACAAAGAAAGAACTTGTACCGAGTTCACTTCCCTGTAGTGGC 3840
 Db 3781 GCTACTGAGTGTGACAAAGAAAGAACTTGTACCGAGTTCACTTCCCTGTAGTGGC 3840
 QY 3841 TCTACCGCTGGCCCAACAGGCTCTGCTGTTGTGATGAGAGAGCGGTGACCGTTTGA 3900
 Db 3841 TCTACCGCTGGCCCAACAGGCTCTGCTGTTGTGATGAGAGAGCGGTGACCGTTTGA 3900
 QY 3901 ATTGAATCTGTGTGGGACCAAGTTCTGCGGTGAGATCAACCCACAGCAAGCTGGATG 3960
 Db 3901 ATTGAATCTGTGTGGGACCAAGTTCTGCGGTGAGATCAACCCACAGCAAGCTGGATG 3960
 QY 3961 GTGGAGGGCCTCTGCTGACATCAGGCTGAGCTGAGCTGAGAGCTGAGAGCTGAGAGCTG 4020
 Db 3961 GTGGAGGGCCTCTGCTGACATCAGGCTGAGCTGAGCTGAGAGCTGAGAGCTGAGAGCTG 4020
 QY 4021 CCTCACTTTGTGGCTCTCAAGGGGGCCATGTGACACATCCCTGTTCCAAATGGGCCAC 4080

Db 4021 CCTCACTTTGTGGCTCTCAAGGGGGCCATGTGGAACATCCCTGTTCCAAATGGGCCAC 4080
 QY 4081 TTTAAAGAGAGGGGATGCTCTCTGAGAAAGCAGCCAGGTGGAGCTCATCACATAGTT 4140
 Db 4081 TTTAAAGAGAGGGGATGCTCTCTGAGAAAGCAGCCAGGTGGAGCTCATCACATAGTT 4140
 QY 4141 CTGGAAAAACCCAGCTTCTCCCTCTGGGAGTCTCTCTGAAAAATGATCATAATGCCCTG 4200
 Db 4141 CTGGAAAAACCCAGCTTCTCCCTCTGGGAGTCTCTCTGAAAAATGATCATAATGCCCTG 4200
 QY 4201 GCTTCATTTCCGTCACCTCTGCTGCTGCTGCTTTTACACCGGCTCATCTGAGGAAGTC 4260
 Db 4201 GCTTCATTTCCGTCACCTCTGCTGCTGCTGCTTTTACACCGGCTCATCTGAGGAAGTC 4260
 QY 4261 ACCTTCACCTCTACCTCATCCCAAGTCACTGCTCATCTCGGAAGAACTGAGCTCTGC 4320
 Db 4261 ACCTTCACCTCTACCTCATCCCAAGTCACTGCTCATCTCGGAAGAACTGAGCTCTGC 4320
 QY 4321 TATCGAAGCCCTGGAGAAAGCCAGCTGTTCTCGGAGTTCTACGTTGGCCACTTTGGATCA 4380
 Db 4321 TATCGAAGCCCTGGAGAAAGCCAGCTGTTCTCGGAGTTCTACGTTGGCCACTTTGGATCA 4380
 QY 4381 GGGATCAGGCTGCAAGTGAAGAAAGAAAGATGAGACTCTGCTGCTGGAGGAGCTTGGTG 4440
 Db 4381 GGGATCAGGCTGCAAGTGAAGAAAGAAAGATGAGACTCTGCTGCTGGAGGAGCTTGGTG 4440
 QY 4441 AAACAGAGAGTCTCATGCTGCACTACTCTGATCCCTCAGCCCGATAGCCGTA 4500
 Db 4441 AAACAGAGAGTCTCATGCTGCACTACTCTGATCCCTCAGCCCGATAGCCGTA 4500
 QY 4501 TCACCTCTGGATGCCCCCAGTGTGCTGCACTTTTGGACCAAGTATCGAGAGCAGCTGATA 4560
 Db 4501 TCACCTCTGGATGCCCCCAGTGTGCTGCACTTTTGGACCAAGTATCGAGAGCAGCTGATA 4560
 QY 4561 GCGGAGTGAATCGGTGGAGGTTGTCTTGGACAACTGCAATGAGCAGGTGTGAGCCAG 4620
 Db 4561 GCGGAGTGAATCGGTGGAGGTTGTCTTGGACAACTGCAATGAGCAGGTGTGAGCCAG 4620
 QY 4621 GAGCAGTACGAGAGGCTGCTGCTGAGAACACGAGGCCAGCCAGATCGGAAGCTGTC 4680
 Db 4621 GAGCAGTACGAGAGGCTGCTGCTGAGAACACGAGGCCAGCCAGATCGGAAGCTGTC 4680
 QY 4681 AGCTTGAGCAGTCTGGGACCGGAAAGTGAAGTGAAGTCTTACCAAGCCCTGAAGGAG 4740
 Db 4681 AGCTTGAGCAGTCTGGGACCGGAAAGTGAAGTGAAGTCTTACCAAGCCCTGAAGGAG 4740
 QY 4741 ACCCATCTCACTCATTTATGGAATCTGGGAGAAAGGAGCAAAAAAGGAGCTCTGCGCA 4800
 Db 4741 ACCCATCTCACTCATTTATGGAATCTGGGAGAAAGGAGCAAAAAAGGAGCTCTGCGCA 4800
 QY 4801 CTCAGCAGCTGAGTATCAACACGAGCCCTTGACCCCTTGAAGCTTGGCTTGGCTGAGCC 4860
 Db 4801 CTCAGCAGCTGAGTATCAACACGAGCCCTTGACCCCTTGAAGCTTGGCTTGGCTGAGCC 4860
 QY 4861 TTTCTTTGGCTCTGAGTTTCTTCTGCAAAACAGTGTGCAATCTGTTTGGCTTCCAGCA 4920
 Db 4861 TTTCTTTGGCTCTGAGTTTCTTCTGCAAAACAGTGTGCAATCTGTTTGGCTTCCAGCA 4920
 QY 4921 CTAAAGTAAATGGAATTTGATGCTTGTCTGGCAATTTGCTCATGCCAGGATG 4980
 Db 4921 CTAAAGTAAATGGAATTTGATGCTTGTCTGGCAATTTGCTCATGCCAGGATG 4980
 QY 4981 CCAGAGGGGGCCCCAGTCCAGTGGCTTAAACAGATCTCAGGAGATGCTCATCTGGAGCT 5040
 Db 4981 CCAGAGGGGGCCCCAGTCCAGTGGCTTAAACAGATCTCAGGAGATGCTCATCTGGAGCT 5040
 QY 5041 GSCAAGACCCCTGACAGCTCATAGAGCTCATCTGCTGGCCACAGCAGCCAGCCTAGA 5100
 Db 5041 GSCAAGACCCCTGACAGCTCATAGAGCTCATCTGCTGGCCACAGCAGCCTAGA 5100
 QY 5101 GCGCTCCGGATCCCATCAGGCGCAAGGAATAGGAGGACATGGAACCATTTGCTCTC 5160

```

Db 5101 GGCCTCCGATCCCATCCAGGCGCAAGAGGAATAGGAGGGAATGAGAACCATTTGCCCTC 5160
QY 5161 TGGCTGTGTACAGAGGTGAGCCCAAAATTTGGGGTTTACGGTGGAGGCCAGTGGATTC 5220
Db 5161 TGGCTGTGTACAGAGGTGAGCCCAAAATTTGGGGTTTACGGTGGAGGCCAGTGGATTC 5220
QY 5221 TTGGCTTTTACAGGAAGATCTACAAGAGCAAGCCAAAGAGTAAAGTGGAAAGGATTTT 5280
Db 5221 TTGGCTTTTACAGGAAGATCTACAAGAGCAAGCCAAAGAGTAAAGTGGAAAGGATTTT 5280
QY 5281 ATTCAAGAAATAAGAGGATACACAGCTCTTTTGAATTTGTACGAGCTTTTCCAGTT 5340
Db 5281 ATTCAAGAAATAAGAGGATACACAGCTCTTTTGAATTTGTACGAGCTTTTCCAGTT 5340
QY 5341 TTTTACAGAAAACCCCTATAAAATTTAAAAATTTTAAATTTTAAAGAAATTTAAAAAAT 5400
Db 5341 TTTTACAGAAAACCCCTATAAAATTTAAAAATTTTAAATTTTAAAGAAATTTAAAAAAT 5400
QY 5401 ACAAAAAGAAAATAAGAAATAAGGAATAAGGAATAAGGAATTAAGCTTACCTAC 5444
Db 5401 ACAAAAAGAAAATAAGAAATAAGGAATAAGGAATAAGGAATTAAGCTTACCTAC 5444

RESULT 3
US-09-956-712-3
; Sequence 3, Application US/09956712
; Publication No. US20030092648A1
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; TITLE OF INVENTION: ANTISENSE MODULATION OF NAC EXPRESSION
; FILE REFERENCE: RTS-0326
; CURRENT APPLICATION NUMBER: US/09/956,712
; CURRENT FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 91
; SEQ ID NO 3
; LENGTH: 5444
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (523)...(4812)
US-09-956-712-3

Query Match 100.0%; Score 5444; DB 11; Length 5444;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCCAGGGCCTGGAGAGGTCTGAAGAAACCTGGGAGCCAGAGCCCGGGGCTCCACTCT 60
Db 1 GCCCCAGGGCCTGGAGAGGTCTGAAGAAACCTGGGAGCCAGAGCCCGGGGCTCCACTCT 60
QY 61 GGGTTCTGAAGCCCATTCCTGCTCTGGGCTCTCCACCCACCTCTTCTCAGCCTT 120
Db 61 GGGTTCTGAAGCCCATTCCTGCTCTGGGCTCTCCACCCACCTCTTCTCAGCCTT 120
QY 121 GCAGCTCAAGGGTTGATCTCAGAGTCCAGGACCCAGGAGGGAAGAAATCTGAGGAA 180
Db 121 GCAGCTCAAGGGTTGATCTCAGAGTCCAGGACCCAGGAGGGAAGAAATCTGAGGAA 180
QY 181 CAGAACAGTGAGGGTTGCCACACCCCATCTCCCGTCAACACATCTCCCTCACCCTCAC 240
Db 181 CAGAACAGTGAGGGTTGCCACACCCCATCTCCCGTCAACACATCTCCCTCACCCTCAC 240
QY 241 CCTCCCTGCTGCGCTTGAGCCCATCCAGGACCTCCCTATCAGCTGACTTCTCCAGT 300
Db 241 CCTCCCTGCTGCGCTTGAGCCCATCCAGGACCTCCCTATCAGCTGACTTCTCCAGT 300
QY 301 GTCTTCGAGGCCCTCTGGGCTCTCCCTCCCTGGCTTTTCTACCACTCCCTCTAT 360
Db 301 GTCTTCGAGGCCCTCTGGGCTCTCCCTCCCTGGCTTTTCTACCACTCCCTCTAT 360
QY 361 CGCGCTCTACTCTGTAGGTGCCCTGGGATTTATAAACTGGGTTCCGAATGCTGAATAAGA 420

```

```

Db 361 CGCGCTCTACTCTGTAGGTGCCCTGGGATTTATAAACTGGGTTCCGAATGCTGAATAAGA 420
QY 421 GACGGTAAAGCAAGGCAAGGACAGCATGTTCTCTGCCTGCTGATACCTCACCAC 480
Db 421 GACGGTAAAGCAAGGCAAGGACAGCATGTTCTCTGCCTGCTGATACCTCACCAC 480
QY 481 CTGGAA CATCCCCAGACACCTCTTAA CTCCGGGACAGAGATGGCTGGGGAGCCTGG 540
Db 481 CTGGAA CATCCCCAGACACCTCTTAA CTCCGGGACAGAGATGGCTGGGGAGCCTGG 540
QY 541 GCGCGCTGGCTGTACTTGGAGTTCCTGAGAGGAGGAGCTGAAAGAGTTCAGCTT 600
Db 541 GCGCGCTGGCTGTACTTGGAGTTCCTGAGAGGAGGAGCTGAAAGAGTTCAGCTT 600
QY 601 CTGCTCGCAATAAAGCGCACTCCAGAGGCTCTTCGGGTGAGACACCCGCTCAGCCAGAG 660
Db 601 CTGCTCGCAATAAAGCGCACTCCAGAGGCTCTTCGGGTGAGACACCCGCTCAGCCAGAG 660
QY 661 AAGCAGTGGCATGGAGTGGCTCTGTA CTTGCTGGCTCAGTATGGGAGCAGCGGGCC 720
Db 661 AAGCAGTGGCATGGAGTGGCTCTGTA CTTGCTGGCTCAGTATGGGAGCAGCGGGCC 720
QY 721 TGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGGCTGAGGTCACTGTGCGCCCAAGCC 780
Db 721 TGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGGCTGAGGTCACTGTGCGCCCAAGCC 780
QY 781 CAGGAAGGGGAGGCGCACTCTCCCTCATTTCCCTACAGCCCAAGTGAACCCCACTGGGG 840
Db 781 CAGGAAGGGGAGGCGCACTCTCCCTCATTTCCCTACAGCCCAAGTGAACCCCACTGGGG 840
QY 841 TCTCCAGCCAAACCCCACTCCACCGCAGTGTCTTAATGSCCTGGATCCATGAATTCGGCGG 900
Db 841 TCTCCAGCCAAACCCCACTCCACCGCAGTGTCTTAATGSCCTGGATCCATGAATTCGGCGG 900
QY 901 GGGTGCAACCGAGGCTCAGAGAGAGGGTTTGAACAAGCTGCTGACACATCTGAGAGC 960
Db 901 GGGTGCAACCGAGGCTCAGAGAGAGGGTTTGAACAAGCTGCTGACACATCTGAGAGC 960
QY 961 CGCTGGAGAGAAATCTCTGCTCTCACTCTCTACCAAGCTCTTCCAAAGTCCCCAGACCAT 1020
Db 961 CGCTGGAGAGAAATCTCTGCTCTCACTCTCTACCAAGCTCTTCCAAAGTCCCCAGACCAT 1020
QY 1021 GAGTCTCCAAGCCAGGAGTCAACCCCAACCCCACTCCACAGAGTGTCTGGGAGCTGG 1080
Db 1021 GAGTCTCCAAGCCAGGAGTCAACCCCAACCCCACTCCACAGAGTGTCTGGGAGCTGG 1080
QY 1081 GGATCCCACTCAGCCAGCCTAGCACCCAGAGAGGAGGCTCTGGGACCCAAATGG 1140
Db 1081 GGATCCCACTCAGCCAGCCTAGCACCCAGAGAGGAGGCTCTGGGACCCAAATGG 1140
QY 1141 CCTCTGGATGAAACCTCAGGAAATTTACTACACAGAAATCAGAGAAAGAGAGAGAGAA 1200
Db 1141 CCTCTGGATGAAACCTCAGGAAATTTACTACACAGAAATCAGAGAAAGAGAGAGAGAA 1200
QY 1201 TCAGAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
Db 1201 TCAGAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
QY 1261 AGCCTACAGCCCAACCA CCA CCAATGGAGCCTTCTGTGAGAGAGAGCCTCTGTTCCACA 1320
Db 1261 AGCCTACAGCCCAACCA CCA CCAATGGAGCCTTCTGTGAGAGAGAGCCTCTGTTCCACA 1320
QY 1321 TGGCCCTGGAAAAATGAGGATTTTAA CCAAAATTTCAACAGCTGCTACTTCTACAAAGA 1380
Db 1321 TGGCCCTGGAAAAATGAGGATTTTAA CCAAAATTTCAACAGCTGCTACTTCTACAAAGA 1380
QY 1381 CCTCAACCCAGAGCAAGATCCCTGGTCAAGAGAGCTGGCCTGATTTATGTGAGGAG 1440
Db 1381 CCTCAACCCAGAGCAAGATCCCTGGTCAAGAGAGCTGGCCTGATTTATGTGAGGAG 1440
QY 1441 AATCAGGAGCAATTTAATTTGAGATCAGAGATTTATTTGGCCCGAGGCTGGATACCCAAAGA 1500

```

Db 1441 AATCGAGGACATTTAAATTGAGATCAGAGACTTATTTTGGCCCGAGGCTGTGATACCAAGAA 1500
Qy 1501 CCTCGCATAGTCACTACTGAGGGGGCTGCTGGAAATTGGGAAGTCAACACTGSCAGGAG 1560
Db 1501 CCTCGCATAGTCACTACTGAGGGGGCTGCTGGAAATTGGGAAGTCAACACTGSCAGGAG 1560
Qy 1561 GTGAAGAAAGCTTGGGGGAGAGCCAGCTGTATATGGGGACCGTTCCAGCATGTCTTCTAC 1620
Db 1561 GTGAAGAAAGCTTGGGGGAGAGCCAGCTGTATATGGGGACCGTTCCAGCATGTCTTCTAC 1620
Qy 1621 TTGAGCTGAGAGAGCTGSCCCAGTCCAAAGGTGGTGAAGTCTGCTGAGCTCATTCGAAAA 1680
Db 1621 TTGAGCTGAGAGAGCTGSCCCAGTCCAAAGGTGGTGAAGTCTGCTGAGCTCATTCGAAAA 1680
Qy 1681 GATGGAGACGCACTCCGGCTCCATTAGACAGATCCTGCTAGGCCAGAGCGGCTGTC 1740
Db 1681 GATGGAGACGCACTCCGGCTCCATTAGACAGATCCTGCTAGGCCAGAGCGGCTGTC 1740
Qy 1741 TTGATCTCGATGGTGTAGATGAGCCAGGATGGGTTCTGAGAGCCGAGTTCTGAGCTC 1800
Db 1741 TTGATCTCGATGGTGTAGATGAGCCAGGATGGGTTCTGAGAGCCGAGTTCTGAGCTC 1800
Qy 1801 TGTCTGCACTGGAGCCAGCCAGCCGCGGATGCACTGCTGGCAGTTTGTCTGGGAAA 1860
Db 1801 TGTCTGCACTGGAGCCAGCCAGCCGCGGATGCACTGCTGGCAGTTTGTCTGGGAAA 1860
Qy 1861 ACTATATCTCCGAGGACCTCTTCCGTGATCAGGCTCGGACACAGCTCTGAGAACCTC 1920
Db 1861 ACTATATCTCCGAGGACCTCTTCCGTGATCAGGCTCGGACACAGCTCTGAGAACCTC 1920
Qy 1921 ATTCTCTCTTTGGAGCAGGACGCTTGGGTAGAGTCTTGGGGTTCTGAGTCCAGCAGG 1980
Db 1921 ATTCTCTCTTTGGAGCAGGACGCTTGGGTAGAGTCTTGGGGTTCTGAGTCCAGCAGG 1980
Qy 1981 AAGGAATAATTTCTACAGATATTTTACAGATGAAAGGCAAGCAATTAGAGCCTTTAGTTG 2040
Db 1981 AAGGAATAATTTCTACAGATATTTTACAGATGAAAGGCAAGCAATTAGAGCCTTTAGTTG 2040
Qy 2041 GTCAATCAAAAGAGCTCTGGGCTGTGTCTTGTGCTGCTGGGTGCTTGGCTGGCC 2100
Db 2041 GTCAATCAAAAGAGCTCTGGGCTGTGTCTTGTGCTGCTGGGTGCTTGGCTGGCC 2100
Qy 2101 TGCACTTGTCTGATGAGCAGATGAAAGGCAAGGAAAACTCACAGTACTTCCAGAGCC 2160
Db 2101 TGCACTTGTCTGATGAGCAGATGAAAGGCAAGGAAAACTCACAGTACTTCCAGAGCC 2160
Qy 2161 ACCACAACCTCTGTCTACATTTACCTTGTCCAGGCTCTCCAGCTCAGCCATTGGAGCC 2220
Db 2161 ACCACAACCTCTGTCTACATTTACCTTGTCCAGGCTCTCCAGCTCAGCCATTGGAGCC 2220
Qy 2221 CAGCTCAGAGACCTCTGCTCTGCTGCTGAGGCACTTGGCAAAAAAGACCTTTTC 2280
Db 2221 CAGCTCAGAGACCTCTGCTCTGCTGCTGAGGCACTTGGCAAAAAAGACCTTTTC 2280
Qy 2281 AGTCCAGATGACCTCAGGAAGCAATGGGTAGATGGGGCCATCATCTCCACCTTCTTGAAG 2340
Db 2281 AGTCCAGATGACCTCAGGAAGCAATGGGTAGATGGGGCCATCATCTCCACCTTCTTGAAG 2340
Qy 2341 ATGGGTATTCTTCAAGAGCACCCCATCTCTGAGCTACAGCTTCAATTCACCTCTGTTTC 2400
Db 2341 ATGGGTATTCTTCAAGAGCACCCCATCTCTGAGCTACAGCTTCAATTCACCTCTGTTTC 2400
Qy 2401 CAAAGGTTCTTTCAGCAATGTCCTATGCTTGGAGGATGAGNAGGGGAGAGGTAAACAT 2460
Db 2401 CAAAGGTTCTTTCAGCAATGTCCTATGCTTGGAGGATGAGNAGGGGAGAGGTAAACAT 2460
Qy 2461 TCTAATTTGATCATAGATTTGAAAAAGACGCTAGAGCATATGGAATACATGGCCCTGTTT 2520
Db 2461 TCTAATTTGATCATAGATTTGAAAAAGACGCTAGAGCATATGGAATACATGGCCCTGTTT 2520
Qy 2521 GGGGCAATCAACCAACGTTTCTTATTTGGGCTGTTTAAAGTATGAGGGGAGAGAGATG 2580
Db 2521 GGGGCAATCAACCAACGTTTCTTATTTGGGCTGTTTAAAGTATGAGGGGAGAGAGATG 2580

Qy 2581 GAGAAATCATCTTTCTACTGCGGCTGTCTCAGGGAGGAACTCTGATGAGTGGTCCCGTCC 2640
Db 2581 GAGAAATCATCTTTCTACTGCGGCTGTCTCAGGGAGGAACTCTGATGAGTGGTCCCGTCC 2640
Qy 2641 CTGCACTGCTGCTGCAAGCCACACTCTCTGAGTCCCTCCACTGCTTGTAGAGACTCGG 2700
Db 2641 CTGCACTGCTGCTGCAAGCCACACTCTCTGAGTCCCTCCACTGCTTGTAGAGACTCGG 2700
Qy 2701 AACAAAAAGTTCTCTGACACAAAGTATGCGCCATTTTCGAGAAATGCGCATGTGTAGAA 2760
Db 2701 AACAAAAAGTTCTCTGACACAAAGTATGCGCCATTTTCGAGAAATGCGCATGTGTAGAA 2760
Qy 2761 ACAGACATGAGAGCTTCTTGTGCACTTTCTGCACTTAAATTCAGCCGCCACGTGAAGAG 2820
Db 2761 ACAGACATGAGAGCTTCTTGTGCACTTTCTGCACTTAAATTCAGCCGCCACGTGAAGAG 2820
Qy 2821 CTTTCACTGATTTGAGGGCAGGACGACAGATCAACATGGAGCCCAACATGGTAGTCTG 2880
Db 2821 CTTTCACTGATTTGAGGGCAGGACGACAGATCAACATGGAGCCCAACATGGTAGTCTG 2880
Qy 2881 TTCAGGTGGTCCCACTGACAGATGCTTATTTGGCAGATTTCTTCTCGTCTCAAGGTC 2940
Db 2881 TTCAGGTGGTCCCACTGACAGATGCTTATTTGGCAGATTTCTTCTCGTCTCAAGGTC 2940
Qy 2941 ACCAGAAAACCTGAAGGAGCTGGACCTAAAGTGAAACTCGCTGAGCCACTCTGCACTGAAG 3000
Db 2941 ACCAGAAAACCTGAAGGAGCTGGACCTAAAGTGAAACTCGCTGAGCCACTCTGCACTGAAG 3000
Qy 3001 AGTCTTTTAAAGACCTCTGAGACGCTCTGCTGCTTGGAGACCTCTGCGGTGCTGCGC 3060
Db 3001 AGTCTTTTAAAGACCTCTGAGACGCTCTGCTGCTTGGAGACCTCTGCGGTGCTGCGC 3060
Qy 3061 TGTGCGCTCAGAGCTGAGGACTGCAAGACCTTGGCTTGGCTGAGAGCAACAGAGCC 3120
Db 3061 TGTGCGCTCAGAGCTGAGGACTGCAAGACCTTGGCTTGGCTGAGAGCAACAGAGCC 3120
Qy 3121 CTGACCGAGCTGGACCTGAGCTTCAATGTGTCTCAGGATGCTGGAGCCAAACACCTTTGC 3180
Db 3121 CTGACCGAGCTGGACCTGAGCTTCAATGTGTCTCAGGATGCTGGAGCCAAACACCTTTGC 3180
Qy 3181 CAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCACTGCACTGCTGCTGCTGCTGCTC 3240
Db 3181 CAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCACTGCACTGCTGCTGCTGCTGCTC 3240
Qy 3241 AGTCTGATGCTGCAAGACCTGCGCTCTGTGCTTGTAGTCCAGCCCGAGCTGAGAGAG 3300
Db 3241 AGTCTGATGCTGCAAGACCTGCGCTCTGTGCTTGTAGTCCAGCCCGAGCTGAGAGAG 3300
Qy 3301 CTGAGACCTGAGCAACAACTGATGACGTTGGCTGCGACTGCTCTGTGAGGGGCTC 3360
Db 3301 CTGAGACCTGAGCAACAACTGATGACGTTGGCTGCGACTGCTCTGTGAGGGGCTC 3360
Qy 3361 AGGCAATCTGCTGCAAACTCATACGCTGGGGCTGGACAGCAAACTCTGAGTGTATGAG 3420
Db 3361 AGGCAATCTGCTGCAAACTCATACGCTGGGGCTGGACAGCAAACTCTGAGTGTATGAG 3420
Qy 3421 ATGAGGAGGAACTGAGGGCCCTGAGCAGGAGAACTCTGAGCTGCTCATCTTACAGAGA 3480
Db 3421 ATGAGGAGGAACTGAGGGCCCTGAGCAGGAGAACTCTGAGCTGCTCATCTTACAGAGA 3480
Qy 3481 CGGAAACCAAGTGTGATGACCCCTACTGAGGCGCTGGATACGGGAGAGATGAGTAATAGC 3540
Db 3481 CGGAAACCAAGTGTGATGACCCCTACTGAGGCGCTGGATACGGGAGAGATGAGTAATAGC 3540
Qy 3541 ACATCTCTCACTCAAGCGGACAGACTCGGATCAGAGAGGGGGCTTCCCATGTTGCTCAG 3600
Db 3541 ACATCTCTCACTCAAGCGGACAGACTCGGATCAGAGAGGGGGCTTCCCATGTTGCTCAG 3600
Qy 3601 GCTAATCTCAAACTCTGAGAGTGAAGATCTTCCCAATTTGCTGAGATTGACAGAGAA 3660
Db 3601 GCTAATCTCAAACTCTGAGAGTGAAGATCTTCCCAATTTGCTGAGATTGACAGAGAA 3660

QY 3661 AGCTCCCGAGAGTAGTACCGGTGGAACTCTTGTGCGTGCCTTCTCTCGCTCTCAAGGG 3720
Db 3661 AGCTCCCGAGAGTAGTACCGGTGGAACTCTTGTGCGTGCCTTCTCTCGCTCTCAAGGG 3720
QY 3721 GACCTGCATACGAAGCCTTTGGGGACTGACGATGACTTCTGGGGCCCCACGGGCGCTG 3780
Db 3721 GACCTGCATACGAAGCCTTTGGGGACTGACGATGACTTCTGGGGCCCCACGGGCGCTG 3780
QY 3781 GCTACTGAGGTAGTTGACAAAGAAAAAGAACTTGTACCGAGTTTCACTTCCCTGTAGCTGGC 3840
Db 3781 GCTACTGAGGTAGTTGACAAAGAAAAAGAACTTGTACCGAGTTTCACTTCCCTGTAGCTGGC 3840
QY 3841 TCCTACCGTGGCCCAACACCGGTCTCTGCTTTGTGATGAGAAAGCGGTGACCGTTGAG 3900
Db 3841 TCCTACCGTGGCCCAACACCGGTCTCTGCTTTGTGATGAGAAAGCGGTGACCGTTGAG 3900
QY 3901 ATTGAATTCGTGTGGGACCAAGTTCTTGGGTGAGATCAACCCACAGCACAGCTGGATG 3960
Db 3901 ATTGAATTCGTGTGGGACCAAGTTCTTGGGTGAGATCAACCCACAGCACAGCTGGATG 3960
QY 3961 GTGGCAGGGCCTCTGCTGGACATCAAGGCTGAGCCTGGAGCTGTGGAAGCTGTGACCTC 4020
Db 3961 GTGGCAGGGCCTCTGCTGGACATCAAGGCTGAGCCTGGAGCTGTGGAAGCTGTGACCTC 4020
QY 4021 CCTCACTTTGTGCTCTTCAAGGGGCCATGTGGAACATCTCTGTTTCCAAATGCCCCAC 4080
Db 4021 CCTCACTTTGTGCTCTTCAAGGGGCCATGTGGAACATCTCTGTTTCCAAATGCCCCAC 4080
QY 4081 TTTAAAGAGGAGGAGTGTCTCTGGAGAGCCAGCCAGGCTGGAGCTGCATCACAAGTT 4140
Db 4081 TTTAAAGAGGAGGAGTGTCTCTGGAGAGCCAGCCAGGCTGGAGCTGCATCACAAGTT 4140
QY 4141 CTGGAACACCCAGCTTCTCCCTTTGGAGTCTCTGAAATGATCATATATGCCCTG 4200
Db 4141 CTGGAACACCCAGCTTCTCCCTTTGGAGTCTCTGAAATGATCATATATGCCCTG 4200
QY 4201 CGCTTCATTCCTGACCTCTGTGCTGTTTACACCGCTCCATCTGAGGAAGTC 4260
Db 4201 CGCTTCATTCCTGACCTCTGTGCTGTTTACACCGCTCCATCTGAGGAAGTC 4260
QY 4261 ACCTTCACCTCTACCTGATCCCAAGTACTGCTTCCATTCGGAAGAACTGGAGCTCTGC 4320
Db 4261 ACCTTCACCTCTACCTGATCCCAAGTACTGCTTCCATTCGGAAGAACTGGAGCTCTGC 4320
QY 4321 TATCGAAGCCCTGGAGAACCAAGCTGTTCTCGAGTTCTACGTTGGCCACTTGGGATCA 4380
Db 4321 TATCGAAGCCCTGGAGAACCAAGCTGTTCTCGAGTTCTACGTTGGCCACTTGGGATCA 4380
QY 4381 GGGATCAGGCTGCAAGTGAAGACAAGATGAGACTCTGTTGGGAGGCTTGGTG 4440
Db 4381 GGGATCAGGCTGCAAGTGAAGACAAGATGAGACTCTGTTGGGAGGCTTGGTG 4440
QY 4441 AAACCGAGGATCTCATGCTGCTGAACTACTCTGATCTCTCCAGCCGCGATAGCGTACCT 4500
Db 4441 AAACCGAGGATCTCATGCTGCTGAACTACTCTGATCTCTCCAGCCGCGATAGCGTACCT 4500
QY 4501 TCACCTCTGATGATCCCGCAGTGTCTGCTGTTGTGACCAAGTATCGAGAGCTGATA 4560
Db 4501 TCACCTCTGATGATCCCGCAGTGTCTGCTGTTGTGACCAAGTATCGAGAGCTGATA 4560
QY 4561 GCCCGAGTGCATCGGTGGAGGTTGCTTGGACAACTGCATGACAGAGTGTGAGCCAG 4620
Db 4561 GCCCGAGTGCATCGGTGGAGGTTGCTTGGACAACTGCATGACAGAGTGTGAGCCAG 4620
QY 4621 GAGCAGTACGAGAGGTTGCTGCTGAGAAACAGAGCCCAAGGATGCGGAAGCTGTTT 4680
Db 4621 GAGCAGTACGAGAGGTTGCTGCTGAGAAACAGAGCCCAAGGATGCGGAAGCTGTTT 4680
QY 4681 AGCTTGAGCCAGTCTCTGGGACCGGAGTGCAGAGTGGACTTACCAAGCCCTGAAGGAG 4740
Db 4681 AGCTTGAGCCAGTCTCTGGGACCGGAGTGCAGAGTGGACTTACCAAGCCCTGAAGGAG 4740
QY 4741 ACCCATCTCACCTTATTGGAACCTCTGGGAGAGGGCAGCAAAAAGGAGCTCTCTGCCA 4800

Db 4741 ACCCATCTCACCTTATTGGAACCTCTGGGAGAGGGCAGCAAAAAGGAGCTCTCTGCCA 4800
QY 4801 CTCACAGCTGAAGTATCAACACACAGCCCTTGACCCCTTGAGTCTCTGGCTTTGGCTGACCC 4860
Db 4801 CTCACAGCTGAAGTATCAACACACAGCCCTTGACCCCTTGAGTCTCTGGCTTTGGCTGACCC 4860
QY 4861 TTTCTTTGGGTCTCAGTCTTTCTCTGCAAAACAAGTTGCCATCTCTGTTTGCCTTCAGCA 4920
Db 4861 TTTCTTTGGGTCTCAGTCTTTCTCTGCAAAACAAGTTGCCATCTCTGTTTGCCTTCAGCA 4920
QY 4921 CTAAAGTAAATGAACTTTGATGATGCTTGTGCGGCATTATGTGTCCATGCGGAGTG 4980
Db 4921 CTAAAGTAAATGAACTTTGATGATGCTTGTGCGGCATTATGTGTCCATGCGGAGTG 4980
QY 4981 CCACAGGGGGCCCCAGTCCAGGTGCCCTAAACAGCATCTCAGGGAATGTCCATCTGGAGCT 5040
Db 4981 CCACAGGGGGCCCCAGTCCAGGTGCCCTAAACAGCATCTCAGGGAATGTCCATCTGGAGCT 5040
QY 5041 GGCAAGACCCCTTGACAGCTCTATAGAGCTCTATCTGTTGGCCACAGCAGCCAAAGCTAGA 5100
Db 5041 GGCAAGACCCCTTGACAGCTCTATAGAGCTCTATCTGTTGGCCACAGCAGCCAAAGCTAGA 5100
QY 5101 GCCCTCCGATCCCATCCAGGCGCAAGAGGAAATAGGAGGACATGGAACCAATTTGCCCTC 5160
Db 5101 GCCCTCCGATCCCATCCAGGCGCAAGAGGAAATAGGAGGACATGGAACCAATTTGCCCTC 5160
QY 5161 TGGCTGTGTACAGGCTGAGCCGCCAAATTTGGGTTTACGCTGGAGGCGCCATGGATTC 5220
Db 5161 TGGCTGTGTACAGGCTGAGCCGCCAAATTTGGGTTTACGCTGGAGGCGCCATGGATTC 5220
QY 5221 TTGGCTTTGTACAGGAAGTCTTCAAGAGCAAGCCAAACAGAGTAAAGTGGAAGGAGTTT 5280
Db 5221 TTGGCTTTGTACAGGAAGTCTTCAAGAGCAAGCCAAACAGAGTAAAGTGGAAGGAGTTT 5280
QY 5281 ATTCAAGAAATTAAGAGGATTCACAGCTCTTTTAGAATTTGTCTAGCAGGCTTTCCAGTT 5340
Db 5281 ATTCAAGAAATTAAGAGGATTCACAGCTCTTTTAGAATTTGTCTAGCAGGCTTTCCAGTT 5340
QY 5341 TTTTACCAGAAAACCCCTATAAATTTAAATTTTACTTTAAATTTAAAGAAATTAAGAAAT 5400
Db 5341 TTTTACCAGAAAACCCCTATAAATTTAAATTTTACTTTAAATTTAAAGAAATTAAGAAAT 5400
QY 5401 ACAAAAAGAAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 5444
Db 5401 ACAAAAAGAAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT 5444

RESULT 4

US-09-956-712-11
; Sequence 11, Application US/09956712
; Publication No. US20030092648A1
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Susan M. Freier
; TITLE OF INVENTION: ANTISENSE MODULATION OF NAC EXPRESSION
; FILE REFERENCE: RTS-0326
; CURRENT APPLICATION NUMBER: US/09/956,712
; CURRENT FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 91
; SEQ ID NO 11
; LENGTH: 6531
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-956-712-11

Query Match 97.1%; Score 5284.2; DB 11; Length 6531;

Best Local Similarity 97.5%; Pred. No. 0;

Matches 5441; Conservative 0; Mismatches 3; Indels 135; Gaps 2;

QY 1 GCCCCAGGCGCTGGAGAGTCTTGAAGAAACCTGGGAGCCAGCAGCCCGGGCTCCACTCT 60

Db	1	GCCCCAGGGCGCTGGAGAGGTCTGAAGAAATCGGGAGCCAGCAGCCCGGGGCTCCACTCT	60
Qy	61	GGGTTCTGAAGCCCAATTCCTGCTCTGGGCTCCTCCACCCACCTCTTCTCAGCCTT	120
Db	61	GGGTTCTGAAGCCCAATTCCTGCTCTGGGCTCCTCCACCCACCTCTTCTCAGCCTT	120
Qy	121	GCAGCTCAAGGGTTGATCTCAGAGTCCAGGACCCAGAGAGGGAAGAATCTGAGGAACA	180
Db	121	GCAGCTCAAGGGTTGATCTCAGAGTCCAGGACCCAGAGAGGGAAGAATCTGAGGAACA	180
Qy	181	CAGAACAGTGAGGTTGCCACACCCCATCTCCCGTCAACAATCTCCCTCACCCCTCAC	240
Db	181	CAGAACAGTGAGGTTGCCACACCCCATCTCCCGTCAACAATCTCCCTCACCCCTCAC	240
Qy	241	CCTCCCTGCTGGCCCTGAGCCCAATCCAGACCTCCCTATCAGCTGACTTCTTCCAGT	300
Db	241	CCTCCCTGCTGGCCCTGAGCCCAATCCAGACCTCCCTATCAGCTGACTTCTTCCAGT	300
Qy	301	GTCTTGAGGCCCTCTGGGCTCCTCCCTCCCTGGCTTTTCTACCACTCCCCCTCTAT	360
Db	301	GTCTTGAGGCCCTCTGGGCTCCTCCCTCCCTGGCTTTTCTACCACTCCCCCTCTAT	360
Qy	361	CGGCGTCTATCTGTAGGTGCCCTGGGATTTATAAACTGGGTTCCGAATGCTGAATAAGA	420
Db	361	CGGCGTCTATCTGTAGGTGCCCTGGGATTTATAAACTGGGTTCCGAATGCTGAATAAGA	420
Qy	421	GACGTTAAGCCCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG	480
Db	421	GACGTTAAGCCCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG	480
Qy	481	CTGGGAAATCCCCAGACACCTCTTAATCCGGGACAGAGATGGCTGGCGAGCCTGG	540
Db	481	CTGGGAAATCCCCAGACACCTCTTAATCCGGGACAGAGATGGCTGGCGAGCCTGG	540
Qy	541	GGCGGCTGGCCTGTATCTTGAAGTTCTTGAAGAGGAGGAGCTGAAGAGTTCCAGCTT	600
Db	541	GGCGGCTGGCCTGTATCTTGAAGTTCTTGAAGAGGAGGAGCTGAAGAGTTCCAGCTT	600
Qy	601	CTGCTGCCAATAAGCGCACTCCAGAGCTCTTGGGTGAGACACCCGCTCAGCCAGAG	660
Db	601	CTGCTGCCAATAAGCGCACTCCAGAGCTCTTGGGTGAGACACCCGCTCAGCCAGAG	660
Qy	661	AAGACAGTGGCATGAGGTGGCTCGTACTTGGTGGCTCAGTATGGGAGCAGCGGGCC	720
Db	661	AAGACAGTGGCATGAGGTGGCTCGTACTTGGTGGCTCAGTATGGGAGCAGCGGGCC	720
Qy	721	TGGGACCTAGCCCTCATCTGGGAGCAGATGGGGCTGAGGTCACTGTGCGCCCAAGCC	780
Db	721	TGGGACCTAGCCCTCATCTGGGAGCAGATGGGGCTGAGGTCACTGTGCGCCCAAGCC	780
Qy	781	CAGGAAGGGGCAAGGCACTCTCCCTCATTTCCCTACAGCCCAAGTGAAACCCACCTGGG	840
Db	781	CAGGAAGGGGCAAGGCACTCTCCCTCATTTCCCTACAGCCCAAGTGAAACCCACCTGGG	840
Qy	841	TCCTCCAGCCAAACCACTCCACCGCAGTGTAAATGCTGGATCCATGAAATGCGGGG	900
Db	841	TCCTCCAGCCAAACCACTCCACCGCAGTGTAAATGCTGGATCCATGAAATGCGGGG	900
Qy	901	GGGTGCAACCCAGGGCTCAGAGAGAGGGTTTGAACAGCTGCTGACACATCTGGAGCC	960
Db	901	GGGTGCAACCCAGGGCTCAGAGAGAGGGTTTGAACAGCTGCTGACACATCTGGAGCC	960
Qy	961	CGCTGGAGAGAAATCTCTGCTCATCTCTTACCAAGCTCTTCCAGCTCCCCAGACCAT	1020
Db	961	CGCTGGAGAGAAATCTCTGCTCATCTCTTACCAAGCTCTTCCAGCTCCCCAGACCAT	1020
Qy	1021	GAGTCTCAGAGGAGTCAACCAAGCCGCCCAATCCACAGCAGCTGCTGGGAGCTGG	1080
Db	1021	GAGTCTCAGAGGAGTCAACCAAGCCGCCCAATCCACAGCAGCTGCTGGGAGCTGG	1080
Qy	1081	GGATCCCAACCTCAGCCAGCCTAGCACCCAGAGCAGGAGCTCTGGGACCCCAATGG	1140
Db	1081	GGATCCCAACCTCAGCCAGCCTAGCACCCAGAGCAGGAGCTCTGGGACCCCAATGG	1140

Qy	1141	CCTCTGGATGAACGCTCAGGAATTTTACTACACAGAAATCAGAGAAAGAGAGAGAGAAA	1200
Db	1141	CCTCTGGATGAACGCTCAGGAATTTTACTACACAGAAATCAGAGAAAGAGAGAGAGAAA	1200
Qy	1201	TCAGAGAAAGGCAAGCCCTCATGGGACGCGTGTAGAAACGCCCCCAACAGSCGACACC	1260
Db	1201	TCAGAGAAAGGCAAGCCCTCATGGGACGCGTGTAGAAACGCCCCCAACAGSCGACACC	1260
Qy	1261	AGCCTACAGCCCCACCAACCCATGGGAGCCTTCTGTGAGAGAGAGCCTCTGTTCACA	1320
Db	1261	AGCCTACAGCCCCACCAACCCATGGGAGCCTTCTGTGAGAGAGAGCCTCTGTTCACA	1320
Qy	1321	TGGCCCTCGAAATAATGAGGATTTTAAACCAAAATTCACACAGCTCTACTTCTACAAA	1380
Db	1321	TGGCCCTCGAAATAATGAGGATTTTAAACCAAAATTCACACAGCTCTACTTCTACAAA	1380
Qy	1381	CCTCAACCCAGAGCAAGATCCCTGTGTCAAGAACTGGCCCTGATTATGTGGAGAG	1440
Db	1381	CCTCAACCCAGAGCAAGATCCCTGTGTCAAGAACTGGCCCTGATTATGTGGAGAG	1440
Qy	1441	AATCAGAGACATTTAATTTGAGATCAGACACTTATTTTGGCCCAAGGCTGGATACCC	1500
Db	1441	AATCAGAGACATTTAATTTGAGATCAGACACTTATTTTGGCCCAAGGCTGGATACCC	1500
Qy	1501	CCTGTCATAGTCATCTCAGGGGCTGCTGCAATTTGGAAAGTCAACACTGGCCAGGCG	1560
Db	1501	CCTGTCATAGTCATCTCAGGGGCTGCTGCAATTTGGAAAGTCAACACTGGCCAGGCG	1560
Qy	1561	GTGAAGGAAGCCTGGGGAGAGGCCAGCTGTATGGGGACCGCTTCACGATGTCTTCTAC	1620
Db	1561	GTGAAGGAAGCCTGGGGAGAGGCCAGCTGTATGGGGACCGCTTCACGATGTCTTCTAC	1620
Qy	1621	TTTACGCTCAGAGAGCTGGCCAGTCAAAGTGTGTGAGTCTGCTGAGCTCATTCGAA	1680
Db	1621	TTTACGCTCAGAGAGCTGGCCAGTCAAAGTGTGTGAGTCTGCTGAGCTCATTCGAA	1680
Qy	1681	GATGGGACAGCCACTCCCGCTCCCAATTAGACAGATCTCTTAGCCAGAGCGGCTGTC	1740
Db	1681	GATGGGACAGCCACTCCCGCTCCCAATTAGACAGATCTCTTAGCCAGAGCGGCTGTC	1740
Qy	1741	TTTATCTCTGATGTGTAGATGAGCCAGATGGGTCTTTCAGGAGCCGAGTCTTCTAG	1800
Db	1741	TTTATCTCTGATGTGTAGATGAGCCAGATGGGTCTTTCAGGAGCCGAGTCTTCTAG	1800
Qy	1801	TGTCTGCACTGAGCCAGCCACAGCCGCGGATGCACTGCTGGGAGTGTCTGGGGAAA	1860
Db	1801	TGTCTGCACTGAGCCAGCCACAGCCGCGGATGCACTGCTGGGAGTGTCTGGGGAAA	1860
Qy	1861	ACTATACCTCCGAGGCACTCTCTGATCAGGCTCGGACCCAGCTCTCAGAACCTC	1920
Db	1861	ACTATACCTCCGAGGCACTCTCTGATCAGGCTCGGACCCAGCTCTCAGAACCTC	1920
Qy	1921	ATTCCTCTCTTGGAGCAGGCACTGTTGGGTAGAGGTCTCTGGGGTCTCTAGTCCAG	1980
Db	1921	ATTCCTCTCTTGGAGCAGGCACTGTTGGGTAGAGGTCTCTGGGGTCTCTAGTCCAG	1980
Qy	1981	AAGGAATATTTCTACAGATATTTCAAGTAAAGGCAAGCAATAGAGCCTTTAGGTTG	2040
Db	1981	AAGGAATATTTCTACAGATATTTCAAGTAAAGGCAAGCAATAGAGCCTTTAGGTTG	2040
Qy	2041	GTCAATCAAAACAAAGAGCTCTGGGCTCTGTGCTTGTGCTGGGTCTCTGGCTGGCC	2100
Db	2041	GTCAATCAAAACAAAGAGCTCTGGGCTCTGTGCTTGTGCTGGGTCTCTGGCTGGCC	2100
Qy	2101	TGCACTTCCTGATGACAGATGAAGCGGAGGAACTCACACTGACTTCCAGACC	2160
Db	2101	TGCACTTCCTGATGACAGATGAAGCGGAGGAACTCACACTGACTTCCAGACC	2160
Qy	2161	ACCAAAACCTCTGTCTACATTAACCTTCCAGGCTCTCCAGCTCAGCCATTTGGAGCC	2220
Db	2161	ACCAAAACCTCTGTCTACATTAACCTTCCAGGCTCTCCAGCTCAGCCATTTGGAGCC	2220

Db 4381 ATGGGCTGTCTTACACTGTGTCTGGGTCTGGTTACGGATGCTGAAATTAATCTCCCAAG 4440
 QY 4306 GAACTGGAGCTCTGCTATCGAAGCCCTGGAGAGACAGCTGTTCTCGGAGTTCTACGTT 4365
 Db 4441 GAACTGGAGCTTTCCTATCGAAGCCCTGGAGAGACAGCTGTTCTCGGAGTTCTACGTT 4500
 QY 4366 GGCACCTTGGGATCAGGGATCAGGCTGCAAGTGAAGACAAAGAAAGATGAGACTCTGGTG 4425
 Db 4501 GGCACCTTGGGATCAGGGATCAGGCTGCAAGTGAAGACAAAGAAAGATGAGACTCTGGTG 4560
 QY 4426 TGGAGGCTTGGTGAACACAGAGATCTCATGCCCTGCAACTACTCTGATTCCTCCAGCC 4485
 Db 4561 TGGAGGCTTGGTGAACACAGAGATCTCATGCCCTGCAACTACTCTGATTCCTCCAGCC 4620
 QY 4486 CGCATAGCCGTACCTTCACTCTGGATGCCCGAGTTGCTGCACTTCTGGACAGTAT 4545
 Db 4621 CGCATAGCCGTACCTTCACTCTGGATGCCCGAGTTGCTGCACTTCTGGACAGTAT 4680
 QY 4546 CGAGAGCAGCTGATAGCCCGAGTGACATCGGTGGAGTTGCTTGGACAAACTGCAATGA 4605
 Db 4681 CGAGAGCAGCTGATAGCCCGAGTGACATCGGTGGAGTTGCTTGGACAAACTGCAATGA 4740
 QY 4606 CAGGTCTGAGCCAGGAGCAGTACGAGAGGTTGCTGGCTGAGAAACAGAGGCCAGCCAG 4665
 Db 4741 CAGGTCTGAGCCAGGAGCAGTACGAGAGGTTGCTGGCTGAGAAACAGAGGCCAGCCAG 4800
 QY 4666 ATGCGAAGCTTTCAGCTTGAACCCAGTCTGGAGCCGGAAGTGCAGATGCACTTAC 4725
 Db 4801 ATGCGAAGCTTTCAGCTTGAACCCAGTCTGGAGCCGGAAGTGCAGATGCACTTAC 4860
 QY 4726 CAAAGCCCTGAAGAGACCCATCTCACTTCAATATGAACTCTGGAGAAAGGCGACAAA 4785
 Db 4861 CAAAGCCCTGAAGAGACCCATCTCACTTCAATATGAACTCTGGAGAAAGGCGACAAA 4920
 QY 4786 AAGGAGCTCTGCACTCAGCAGCTGAAGTATCAACACAGCCCTTTCAGCCCTTGAAGTCT 4845
 Db 4921 AAGGAGCTCTGCACTCAGCAGCTGAAGTATCAACACAGCCCTTTCAGCCCTTGAAGTCT 4980
 QY 4846 GGGTTTGGCTGACCTTCTTGGGTCTCAGTTTCTTCTTCTGCAACCAAGTTGCCATCTG 4905
 Db 4981 GGGTTTGGCTGACCTTCTTGGGTCTCAGTTTCTTCTTCTGCAACCAAGTTGCCATCTG 5040
 QY 4906 GTTGTGCTTCCAGCACTAAAGTAAATGGAACTTTGATGATGCCCTTTCCTGGGCATATGTG 4965
 Db 5041 GTTGTGCTTCCAGCACTAAAGTAAATGGAACTTTGATGATGCCCTTTCCTGGGCATATGTG 5100
 QY 4966 TCCATGCCAGGATGCCACAGGGGGCCCAAGTCCAGTGGCTTAAACAGCATCTCAGGGAA 5025
 Db 5101 TCCATGCCAGGATGCCACAGGGGGCCCAAGTCCAGTGGCTTAAACAGCATCTCAGGGAA 5160
 QY 5026 TGTCCATCTGGAGCTGGCAAGCCCTGAGACCTCATAGAGCTCATCTGGTGGCCACA 5085
 Db 5161 TGTCCATCTGGAGCTGGCAAGCCCTGAGACCTCATAGAGCTCATCTGGTGGCCACA 5220
 QY 5086 GCAGCCAAAGCTAGAGCTTCCGATCCATCCAGGCGCAAGAGAAATAGAGGGACAT 5145
 Db 5221 GCAGCCAAAGCTAGAGCTTCCGATCCATCCAGGCGCAAGAGAAATAGAGGGACAT 5280
 QY 5146 GGAACCATTTGCTCTGGCTGTGTACAGGGTGGAGCCCAAAATTCGGGTTCAGCGTGG 5205
 Db 5281 GGAACCATTTGCTCTGGCTGTGTACAGGGTGGAGCCCAAAATTCGGGTTCAGCGTGG 5340
 QY 5206 AGGCCAAGTGGATCTTGGCTTGTGTACAGGAATCTACAGAGCAAGCAACAGAGTAA 5265
 Db 5341 AGGCCAAGTGGATCTTGGCTTGTGTACAGGAATCTACAGAGCAAGCAACAGAGTAA 5400
 QY 5266 AGTGGAAAGAGTTTATTCAGAAAAATAAGGAGTATCAGAGCTCTTTTATAGAAATTCCTTA 5325
 Db 5401 AGTGGAAAGAGTTTATTCAGAAAAATAAGGAGTATCAGAGCTCTTTTATAGAAATTCCTTA 5460
 QY 5326 GCAGGCTTCCAGTTTATACAGAAACCCCTATAAATTAATAATTTTACTTAAATTT 5385
 Db 5461 GCAGGCTTCCAGTTTATACAGAAACCCCTATAAATTAATAATTTTACTTAAATTT 5520

QY 5386 AAGAATTAAAAAATACAAAAAAGAAAAATGAAAAATAAGGAATAAGAAAGTTACTTAC 5444
 Db 5521 AAGAATTAAAAAATACAAAAAAGAAAAATGAAAAATAAGGAATAAGAAAGTTACTTAC 5579

RESULT 5
 US-09-956-712-12
 ; Sequence 12, Application US/09956712
 ; Publication No. US20030092648A1
 ; GENERAL INFORMATION:
 ; APPLICANT: C. Frank Bennett
 ; APPLICANT: Susan M. Freier
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF NAC EXPRESSION
 ; FILE REFERENCE: RTS-0326
 ; CURRENT APPLICATION NUMBER: US/09/956,712
 ; CURRENT FILING DATE: 2001-09-19
 ; NUMBER OF SEQ ID NOS: 91
 ; SEQ ID NO 12
 ; LENGTH: 5100
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (160) ... (4449)
 US-09-956-712-12

Query Match 89.7%; Score 4882.4; DB 11; Length 5100;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 4898; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 516 GACAGAGATGGCTGGGAGCCTGGGCGCTGGCTGTTACTTTGGAGTTCTCTCAAGAA 575
 Db 153 GAACAGATGGCTGGGAGCCTGGGCGCTGGCTGTTACTTTGGAGTTCTCTCAAGAA 212
 QY 576 GGAGGAGCTGAAGGAGTTCTCAGCTTCTGCTGCCAATAAAGCGCACTCCAGAGCTTTC 635
 Db 213 GGAGGAGCTGAAGGAGTTCTCAGCTTCTGCTGCCAATAAAGCGCACTCCAGAGCTTTC 272
 QY 636 GGGTGAAGACCCGCTCAGCCAGAGACAGAGTGGCATGGAGTGGCTCTGACTCTGT 695
 Db 273 GGGTGAAGACCCGCTCAGCCAGAGACAGAGTGGCATGGAGTGGCTCTGACTCTGT 332
 QY 696 GGCTCAGTATGGGAGCAGCGGCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGG 755
 Db 333 GGCTCAGTATGGGAGCAGCGGCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGG 392
 QY 756 GCTGAGGTCACTGTGCGCCCAAGCCAGGAAAGGGGAGGCCCACTCTCCTCAATCCCTTA 815
 Db 393 GCTGAGGTCACTGTGCGCCCAAGCCAGGAAAGGGGAGGCCCACTCTCCTCAATCCCTTA 452
 QY 816 CAGCCCAAGTGAACCCCACTGGGGTCTCCAGCCAAACCCACCTCCAGCGAGTCTAAT 875
 Db 453 CAGCCCAAGTGAACCCCACTGGGGTCTCCAGCCAAACCCACCTCCAGCGAGTCTAAT 512
 QY 876 GCGCTGGATCCATGAATTTGCGCGGGGTGCAACCCAGGCTCAGAGAGAGGGTTTGAG 935
 Db 513 GCGCTGGATCCATGAATTTGCGCGGGGTGCAACCCAGGCTCAGAGAGAGGGTTTGAG 572
 QY 936 ACAGTGTGCTGACACATCTGGAAGCGCTGGAGAGAAATCTTGCTCACTCTCTACCA 995
 Db 573 ACAGTGTGCTGACACATCTGGAAGCGCTGGAGAGAAATCTTGCTCACTCTACCA 632
 QY 996 AGCTCTTCCAGCTCCCGAGCAGTCTGCTTCCAGAGGAGTCAACCAAGCCGCCAC 1055
 Db 633 AGCTCTTCCAGCTCCCGAGCAGTCTGCTTCCAGAGGAGTCAACCAAGCCGCCAC 692
 QY 1056 ATCCACAGCAGTCTGGGAGCTGGGATCCCACTCAGCCAGCTAGCAGACCCAGAGA 1115
 Db 693 ATCCACAGCAGTCTGGGAGCTGGGATCCCACTCAGCCAGCTAGCAGACCCAGAGA 752
 QY 1116 GCAGGAGGCTCCTGGGAGCCCAATGGCTCTGGATGAACGTCAGGAAATTTACTACACAGA 1175

Db 753 GGAGGAGGCTCTCGGACCAATGGCCTCTGATGAACGTCAGGAATTTACTACACAGA 812
Qy 1176 AATCAGAGAAAGAGAGAGAGAGAAATCAGAGAAAGCGAGGCCCCCATGGCGAGCGTGTG 1235
Db 813 AATCAGAGAAAGAGAGAGAGAGAAATCAGAGAAAGCGAGGCCCCCATGGCGAGCGTGTG 872
Qy 1236 AGGAACGCCCCCAAGGCGCACAACAGCCTACAGCCCAACACACCCATGGAGGAGCCTTC 1295
Db 873 AGGAACGCCCCCAAGGCGCACAACAGCCTACAGCCCAACACACCCATGGAGGAGCCTTC 932
Qy 1296 TGTGAGAGAGAGCCTCTGTCTTCCACATGGCCCTGGAAATAGAGATTTTAAACCAAAATTT 1355
Db 933 TGTGAGAGAGAGCCTCTGTCTTCCACATGGCCCTGGAAATAGAGATTTTAAACCAAAATTT 992
Qy 1356 CACACAGCTGTACTTTCTACAAAGA CTTACACCCCAAGAACCAAGATCCCTGGTCAAAG 1415
Db 993 CACACAGCTGTACTTTCTACAAAGACCTCACCCCAAGAACCAAGATCCCTGGTCAAAG 1052
Qy 1416 AAGCTGGCCTGATTTATGTGGAGAGAAATCGAGGACATTTAAATGAGATCAGAGACTTAT 1475
Db 1053 AAGCTGGCCTGATTTATGTGGAGAGAAATCGAGGACATTTAAATGAGATCAGAGACTTAT 1112
Qy 1476 TGGCCCAAGCCTGGATACCCAAAGACCTCGCATAGTCACTACGAGGGGCTGCTGGAAT 1535
Db 1113 TGGCCCAAGCCTGGATACCCAAAGACCTCGCATAGTCACTACGAGGGGCTGCTGGAAT 1172
Qy 1536 TGGGAAGTCAACA CTTGGCCAGGCGAGTGAAGAACCTCGGGGAGAGGCCAGCTGTATGG 1595
Db 1173 TGGGAAGTCAACA CTTGGCCAGGCGAGTGAAGAACCTCGGGGAGAGGCCAGCTGTATGG 1232
Qy 1596 GGAACCGCTCCAGCATGTCTTCTACTTACGCTGACAGAGCTGGCCAGTCCAAAGTGTG 1655
Db 1233 GGAACCGCTCCAGCATGTCTTCTACTTACGCTGACAGAGCTGGCCAGTCCAAAGTGTG 1292
Qy 1656 GAGTCTCGCTGAGCTCATCGGAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGAT 1715
Db 1293 GAGTCTCGCTGAGCTCATCGGAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGAT 1352
Qy 1716 CTTGCTTAGGCGAGAGCGGCTGTCTTCTATCTTCTGATGTGTAGATGAGCCAGGATGGGT 1775
Db 1353 CTTGCTTAGGCGAGAGCGGCTGTCTTCTATCTTCTGATGTGTAGATGAGCCAGGATGGGT 1412
Qy 1776 CTTGACGAGCCGAGTCTGAGCTGTCTGTGCACTGGAGCCAGCCACAGCCCGCGGATGC 1835
Db 1413 CTTGACGAGCCGAGTCTGAGCTGTCTGTGCACTGGAGCCAGCCACAGCCCGCGGATGC 1472
Qy 1836 ACTGCTGGGAGTGTCTGGGGAACCTATACTTCCCGAGGATCCTTCTGATCAAGGC 1895
Db 1473 ACTGCTGGGAGTGTCTGGGGAACCTATACTTCCCGAGGATCCTTCTGATCAAGGC 1532
Qy 1896 TCGGACACAGCTCTGACAGACTCATTTCTTCTTTGGAGCAGGCACGTTGGGTAGAGGT 1955
Db 1533 TCGGACACAGCTCTGACAGACTCATTTCTTCTTTGGAGCAGGCACGTTGGGTAGAGGT 1592
Qy 1956 CTTGGGGTCTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTTACAGATGAAG 2015
Db 1593 CTTGGGGTCTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTTACAGATGAAG 1652
Qy 2016 GCAGCAATTAGAGCCTTTAGTGTGTGTCAAA TCAAA CAAAGAGCTCTGGGCCCTGTGTCT 2075
Db 1653 GCAGCAATTAGAGCCTTTAGTGTGTGTCAAA TCAAA CAAAGAGCTCTGGGCCCTGTGTCT 1712
Qy 2076 TGTGCCCTGGGTCTCGGCTGGCTGCACTTCCCTCTTTGGAGCAGGCACGTTGGGTAGAGGT 2135
Db 1713 TGTGCCCTGGGTCTCGGCTGGCTGCACTTCCCTCTTTGGAGCAGGCACGTTGGGTAGAGGT 1772
Qy 2136 AAAA CTTACACTGACTTCCAAAGACCAACCA CCAACCCCTCTGTCTACATTA CTTGCCCAGGC 2195
Db 1773 AAAA CTTACACTGACTTCCAAAGACCAACCA CCAACCCCTCTGTCTACATTA CTTGCCCAGGC 1832
Qy 2196 TCTCCAAGCTCAGCCATTTGGGACCCAGCTCAGAGACCTTGTCTCTGTGCTGCTGAGG 2255
Db 1833 TCTCCAAGCTCAGCCATTTGGGACCCAGCTCAGAGACCTTGTCTCTGTGCTGCTGAGG 1892

Qy 2256 CATCTGGCAAAAAAGAACCCCTTTTTCAGTCCAGATCACTCAGGAAGCATGGGTATTAGATGG 2315
Db 1893 CATCTGGCAAAAAAGAACCCCTTTTTCAGTCCAGATCACTCAGGAAGCATGGGTATTAGATGG 1952
Qy 2316 GGCATCATCTCCACCTCTTTGAAGATGGGTATTCTTCAAGAGCAACCCCATCCCTCTGAG 2375
Db 1953 GGCATCATCTCCACCTCTTTGAAGATGGGTATTCTTCAAGAGCAACCCCATCCCTCTGAG 2012
Qy 2376 CTACAGCTTCATTTCACCTCTGTTCCTCAAGAGTCTTTTGCAGCAATGTCTATGTCTTGA 2435
Db 2013 CTACAGCTTCATTTCACCTCTGTTCCTCAAGAGTCTTTTGCAGCAATGTCTATGTCTTGA 2072
Qy 2436 GGATCAGAGGGGAGAGAGGTAAACATTTCTAAATTTGCATCATAGATTTTGGAAAAAGOCGTAGA 2495
Db 2073 GGATCAGAGGGGAGAGAGGTAAACATTTCTAAATTTGCATCATAGATTTTGGAAAAAGOCGTAGA 2132
Qy 2496 AGCATATGGAAATACATGGCCTGTTCCTGGGCATCAACACACACGTTTCTTATTTGGGCCCTGT 2555
Db 2133 AGCATATGGAAATACATGGCCTGTTCCTGGGCATCAACACACACGTTTCTTATTTGGGCCCTGT 2192
Qy 2556 AAGTATGAGGGGAGAGAGATGGAGAAACATCTTTTCACTCCCGGCTGTCTCAGGGGAG 2615
Db 2193 AAGTATGAGGGGAGAGAGATGGAGAAACATCTTTTCACTCCCGGCTGTCTCAGGGGAG 2252
Qy 2616 GAACTGATGACAGTGGGTCCGCTCCCTGACGTGCTGTGACGCCACACTCTCTGGAGTC 2675
Db 2253 GAACTGATGACAGTGGGTCCGCTCCCTGACGTGCTGTGACGCCACACTCTCTGGAGTC 2312
Qy 2676 CTTCCACTGCTGTATGACGACTCGGAAACAAACGTTTCTTGAACACAGTGTATGGCCCATTT 2735
Db 2313 CTTCCACTGCTGTATGACGACTCGGAAACAAACGTTTCTTGAACACAGTGTATGGCCCATTT 2372
Qy 2736 CGAAGAAATGGSCATGTGTGTAGAAA CAGACATGAGACTCTTAGTGTGACATTTCTGCAAT 2795
Db 2373 CGAAGAAATGGSCATGTGTGTAGAAA CAGACATGAGACTCTTAGTGTGACATTTCTGCAAT 2432
Qy 2796 TAAATTCAGCCGCCACGCTGAAGAACTTCAGCTGATTTGAGGGCAGGCAGCAGATCAAC 2855
Db 2433 TAAATTCAGCCGCCACGCTGAAGAACTTCAGCTGATTTGAGGGCAGGCAGCAGATCAAC 2492
Qy 2856 ATGGAGCCCCACATGGTGTAGTCTCTGTTAGTGGGTCCAGTCCAGTACAGATGCCTATTGGCA 2915
Db 2493 ATGGAGCCCCACATGGTGTAGTCTCTGTTAGTGGGTCCAGTCCAGTACAGATGCCTATTGGCA 2552
Qy 2916 GATTCTTCTCCGTCCTCAAGGTCACAGGTCACCAAGAACTGAGGAGCTGGAACCTTAAGTGAAG 2975
Db 2553 GATTCTTCTCCGTCCTCAAGGTCACCAAGAACTGAGGAGCTGGAACCTTAAGTGAAG 2612
Qy 2976 CTCGCTGAGCCCACTCTGACAGTGAAGAGTCTTTTGAAGACCTTGAAGACCCCTCGCTGCCT 3035
Db 2613 CTCGCTGAGCCCACTCTGACAGTGAAGAGTCTTTTGAAGACCCCTCGCTGCCT 2672
Qy 3036 CTTGAGAGCCCTGCGGTGGCTGGCTGTGGCTCTACAGCTGAGGACTGCAAGGACCTTGC 3095
Db 2673 CTTGAGAGCCCTGCGGTGGCTGGCTGTGGCTCTACAGCTGAGGACTGTAAGGACCTTGC 2732
Qy 3096 CTTTGGGCTGAGAGCCCAACCCAGACCCCTGACCGAGCTGGACCTTCAATGTGTCTAC 3155
Db 2733 CTTTGGGCTGAGAGCCCAACCCAGACCCCTGACCGAGCTGGACCTTCAATGTGTCTAC 2792
Qy 3156 GGATGCTGGAGCCAAACACCTTTTGCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCG 3215
Db 2793 GGATGCTGGAGCCAAACACCTTTTGCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCG 2852
Qy 3216 ACTGACAGCTGGTCAAGTGTGGCTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3275
Db 2853 ACTGACAGCTGGTCAAGTGTGGCTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2912
Qy 3276 TAGTGCAGCCCCAGCCTGAAGGAGCTAGACCTTGACGAGAAACCTTGATGACGTTGG 3335
Db 2913 TAGTGCAGCCCCAGCCTGAAGGAGCTAGACCTTGACGAGAAACCTTGATGACGTTGG 2972

QY 3336 CGTGGAGCTGCTCTGTGAGGGGCTCAGGCATCCTGCTGCAAACTCATACGCTGGGGCT 3395
 Db 2973 CGTGGAGCTGCTCTGTGAGGGGCTCAGGCATCCTGCTGCAAACTCATACGCTGGGGCT 3032
 QY 3396 GGACACAGAACTCTGAGTGATGAGATGAGGAGGAACTGAGGGCCCTGGAGCAGAGAA 3455
 Db 3033 GGACACAGAGCTCTGAGTGATGAGATGAGGAGGAACTGAGGGCCCTGGAGCAGAGAA 3092
 QY 3456 AACTGAGCTGCTCATCTTCAGCAGACGGAAACCAAGTGTGATGACCCCTACTGAGGGCT 3515
 Db 3093 GCTTCAGCTGCTCATCTTCAGCAGACGGAAACCAAGTGTGATGACCCCTACTGAGGGCT 3152
 QY 3516 GGATACGGAGAGATGAGTAATAGCACATCCTCACCTCAAGCGGCAGAGACTCGGATCAGA 3575
 Db 3153 GGATACGGAGAGATGAGTAATAGCACATCCTCACCTCAAGCGGCAGAGACTCGGATCAGA 3212
 QY 3576 GAGGGCGGCTTCCATGTTGCTCAGGCTAATCTCAAACTCCTGGAAGTGGAGTGGT 3635
 Db 3213 GAGGGCGGCTTCCATGTTGCTCAGGCTAATCTCAAACTCCTGGAAGTGGAGTGGT 3272
 QY 3636 CCCAATTGCTGAGATTGCAAGGAAAGCTCCAGAGGTAGTACCGGTGGAATCTTGTG 3695
 Db 3273 CCCAATTGCTGAGATTGCAAGGAAAGCTCCAGAGGTAGTACCGGTGGAATCTTGTG 3332
 QY 3696 CGTGCCTTCTCCTGCTCTCAAGGGACCTGCATACGAAGCCCTTTGGGACTGACGATGA 3755
 Db 3333 CGTGCCTTCTCCTGCTCTCAAGGGACCTGCATACGAAGCCCTTTGGGACTGACGATGA 3392
 QY 3756 CTTTCTGGGGCCCAACGGGGCTGTGCTACTGAGGTAGTGAACAAAGAAAGAACTTGTGA 3815
 Db 3393 CTTTCTGGGGCCCAACGGGGCTGTGCTACTGAGGTAGTGAACAAAGAAAGAACTTGTGA 3452
 QY 3816 CCGAGTTCACTTCCCTGTAGCTGGCTTACCGCTGGCCCAACCGGCTCTGCTTGT 3875
 Db 3453 CCGAGTTCACTTCCCTGTAGCTGGCTTACCGCTGGCCCAACCGGCTCTGCTTGT 3512
 QY 3876 GATGAGAGAGCGGTGACCGTTGAGATTGAAATCTGTGTGGGACAGTTCCTGGGTGA 3935
 Db 3513 GGTGAGAGAGCGGTGACCGTTGAGATTGAAATCTGTGTGGGACAGTTCCTGGGTGA 3572
 QY 3936 GATCAACCCACAGCAGCTGGATGTGGCAGGGCTCTGCTGGAATCAAGCTGAGCC 3995
 Db 3573 GATCAACCCACAGCAGCTGGATGTGGCAGGGCTCTGCTGGAATCAAGCTGAGCC 3632
 QY 3996 TGGAGCTGTGAGAGCTGTGCACTTCCCTCAGTGTGGCTCTCAAGGGGCCATGTGA 4055
 Db 3633 TGGAGCTGTGAGAGCTGTGCACTTCCCTCAGTGTGGCTCTCAAGGGGCCATGTGA 3692
 QY 4056 CACATCCCTGTTCCAAATGGGCCACTTAAAGAGAGGGGATGCTCCTGGAGAGCCAGC 4115
 Db 3693 CACATCCCTGTTCCAAATGGGCCACTTAAAGAGAGGGGATGCTCCTGGAGAGCCAGC 3752
 QY 4116 CAGGGTGGAGCTGCATCACAATGTTCTGGAAAAACCCAGCTTCTCCCCCTGGAGTCTC 4175
 Db 3753 CAGGGTGGAGCTGCATCACAATGTTCTGGAAAAACCCAGCTTCTCCCCCTGGAGTCTC 3812
 QY 4176 CTTGAAAAATGATCCATAATGCTGCTGCTTCAATCCCGTCACTCTGTGGTGTGCTTTA 4235
 Db 3813 CTTGAAAAATGATCCATAATGCTGCTGCTTCAATCCCGTCACTCTGTGGTGTGCTTTA 3872
 QY 4236 CCACCGCTCCATCTCTGAGGAGTCACTTCCACCTCTACCTGATCCCAAGTGAATGCTC 4295
 Db 3873 CCACCGCTCCATCTCTGAGGAGTCACTTCCACCTCTACCTGATCCCAAGTGAATGCTC 3932
 QY 4296 CATTCGGAAAGGAATGGAGCTGTGCTATCGAAGCCCTGGAGAGACCAAGCTGTCTCGGA 4355
 Db 3933 CATTCGGAAAGGAATGGAGCTGTGCTATCGAAGCCCTGGAGAGACCAAGCTGTCTCGGA 3992
 QY 4356 GTTCTACGTTGGCCACTTGGGATCAGGGATCAGGCTGCAAGTGAAGAAACAGAAAGATGA 4415
 Db 3993 GTTCTACGTTGGCCACTTGGGATCAGGGATCAGGCTGCAAGTGAAGAAACAGAAAGATGA 4052
 QY 4416 GACTCTGGTGTGGAGGGCTTGGTGAAACCAAGGAGATCTCATGCCCTGCAACTACTCTGAT 4475

RESULT 6

4053 GACTCTGCTGTGGAGGGCTTGGTGAAACCAAGAGATCTCATGCCCTGCAACTACTCTGAT 4112
 QY 4476 CCCTCCAGCCCGCATAGCCGTACCTTCACTCTGATGCCCCGCGAGTTGCTGCACTTGT 4535
 Db 4113 CCCTCCAGCCCGCATAGCCGTACCTTCACTCTGATGCCCCGCGAGTTGCTGCACTTGT 4172
 QY 4536 GGACAGTATCGAGAGCAGCTGATAGCCCGAGTGCATCGGTGGAGTTGTCTTGGACAA 4595
 Db 4173 GGACAGTATCGAGAGCAGCTGATAGCCCGAGTGCATCGGTGGAGTTGTCTTGGACAA 4232
 QY 4596 ACTGCATGACAGGTGCTGAGCCAGGACGATPACGAGAGGTGCTGGCTGAGAACACGAG 4655
 Db 4233 ACTGCATGACAGGTGCTGAGCCAGGACGATPACGAGAGGTGCTGGCTGAGAACACGAG 4292
 QY 4656 GCCCAGCCAGATGCCGAGAGCTGTTGAGCTTGAGCCAGTCTCTGGGACCGGAGTGCAGAA 4715
 Db 4293 GCCCAGCCAGATGCCGAGAGCTGTTGAGCTTGAGCCAGTCTCTGGGACCGGAGTGCAGAA 4352
 QY 4716 TGGACTCTACCAAGCCCTGAGGAGACCATCCTCACTCATTTATGGAACCTCTGGAGAA 4775
 Db 4353 TGGACTCTACCAAGCCCTGAGGAGACCATCCTCACTCATTTATGGAACCTCTGGAGAA 4412
 QY 4776 GGGCAGCAAAAAGGAGCTCTGCCACTCAGCAGCTGAAAGTATCAAACACAGCCCTTGACC 4835
 Db 4413 GGGCAGCAAAAAGGAGCTCTGCCACTCAGCAGCTGAAAGTATCAAACACAGCCCTTGACC 4472
 QY 4836 CTTGAGTCTGGCTTTGGCTGACCCCTCTTGGGTCTCAGTTTCTCTCTGCAAAACAG 4895
 Db 4473 CTTGAGTCTGGCTTTGGCTGACCCCTCTTGGGTCTCAGTTTCTCTCTGCAAAACAG 4532
 QY 4896 TTGCCATCTGGTTTGCCTTCCAGCACTAAAGTAAATGGAATCTTGTATGATGCTTGTCTGG 4955
 Db 4533 TTGCCATCTGGTTTGCCTTCCAGCACTAAAGTAAATGGAATCTTGTATGATGCTTGTCTGG 4592
 QY 4956 GCATTTATGTGTCCATGCCCCAGGAGTCCACAGGGGCCCCAGTCCAGGTGGCCTAAACAG 5015
 Db 4593 GCATTTATGTGTCCATGCCCCAGGAGTCCACAGGGGCCCCAGTCCAGGTGGCCTAAACAG 4652
 QY 5016 TCTCAGGGAATGTCCATCTGGAGCTGGCAGACCCCTCGAGACCTCATAGAGCCCTCATCT 5075
 Db 4653 TCTCAGGGAATGTCCATCTGGAGCTGGCAGACCCCTCGAGACCTCATAGAGCCCTCATCT 4712
 QY 5076 GGTGCCCCACAGCAGCAAGCCTTAGAGCCCTCCGATCCCATCCAGGCGCAAGAGAGAAATA 5135
 Db 4713 GGTGCCCCACAGCAGCAAGCCTTAGAGCCCTCCGATCCCATCCAGGCGCAAGAGAGAAATA 4772
 QY 5136 GGAGGGACATGGAAACCATTTGCTCTGCTGTGTCAAGGGTGAAGCCCAAAATTTGGGT 5195
 Db 4773 GGAGGGACATGGAAACCATTTGCTCTGCTGTGTCAAGGGTGAAGCCCAAAATTTGGGT 4832
 QY 5196 TCAGCGTGGAGGCCACAGTGGATTTCTTGGCTTTGTACAGGAAGATCTACAGAGCAAGCC 5255
 Db 4833 TCAGCGTGGAGGCCACAGTGGATTTCTTGGCTTTGTACAGGAAGATCTACAGAGCAAGCC 4892
 QY 5256 AACACAGTAAAGTGGAAAGGAGTTTATTCAGAAAAATAAGAGAGTATCAAGAGTCTTTTAG 5315
 Db 4893 AACACAGTAAAGTGGAAAGGAGTTTATTCAGAAAAATAAGAGAGTATCAAGAGTCTTTTAG 4952
 QY 5316 AATTTTGTCTAGCAGGCTTTCCAGTTTTCACAGAAAAACCCCTATAAAATTTAAAAATTTT 5375
 Db 4953 AATTTTGTCTAGCAGGCTTTCCAGTTTTCACAGAAAAACCCCTATAAAATTTAAAAATTTT 5012
 QY 5376 ACTTAAATTTAAGAAATTTAAAAATAACAAAAAGAAAAATGAAAAATAAGAAATAAGAA 5435
 Db 5013 ACTTAAATTTAAGAAATTTAAAAATAACAAAAAGAAAAATGAAAAATAAGAAATAAGAA 5072
 QY 5436 GTTA 5439
 Db 5073 GTTA 5076

661	GAAAGAGAGAGAGAAATCAGAGAAAGCAGGCCCCCATGGGCAAGCGTGGTAGGAACG	720
11243	CCCCCAGGGCGCACACGAGCCTACAGCCCCACCAACCCATGGAGCCTTCTGTGAGA	1302
721	CCCCCAGGGCGCACACGAGCCTACAGCCCCACCAACCCATGGAGCCTTCTGTGAGA	780
13303	GAGAGCCTCTGTTTCCATGATGGCCCTGGAAAAATGAGGATTTTAAACCAAAAATTCACACAG	1362
781	GAGAGCCTCTGTTTCCATGATGGCCCTGGAAAAATGAGGATTTTAAACCAAAAATTCACACAG	840
1363	CTGCTACTTCTCAAAAGACTCACCCCAAGCAGAGATCCCTCGTCAAGAGAAAGCTGG	1422
841	CTGCTACTTCTCAAAAGACTCACCCCAAGCAGAGATCCCTCGTCAAGAGAAAGCTGG	900
1423	CCTGATATGTGGAGGAGAAATCGAGGACATTTAAATTGAGATCAGAGACTTATTTTGGCCCA	1482
901	CCTGATATGTGGAGGAGAAATCGAGGACATTTAAATTGAGATCAGAGACTTATTTTGGCCCA	960
1483	GGCTGGATACCAAGAACTCTGCATAGTCACTATGCAAGGGGCTGCTGGAAATTTGGGAAG	1542
961	GGCTGGATACCAAGAACTCTGCATAGTCACTATGCAAGGGGCTGCTGGAAATTTGGGAAG	1020
1543	TCAACCTGSCAGGACAGGTGAAGGAGCCTCGGGGAGAGGCCAGCTGATGGGACCGC	1602
11021	TCAACCTGSCAGGACAGGTGAAGGAGCCTCGGGGAGAGGCCAGCTGATGGGACCGC	1080
1603	TTCCAGCATCTCTTCTACTTTACGTCGAGAGAGCTGGCCCACTCCAAGGTGGTAGTCTC	1662
11081	TTCCAGCATCTCTTCTACTTTACGTCGAGAGAGCTGGCCCACTCCAAGGTGGTAGTCTC	1140
1663	GCTGAGCTCATTCGAAAAAGATGGACAGCCACTCCGGCTCCCATTTAGACAGATCCTGTCT	1722
11141	GCTGAGCTCATTCGAAAAAGATGGACAGCCACTCCGGCTCCCATTTAGACAGATCCTGTCT	1200
1723	AGGCCAGAGCGGCTGCTTCACTCCGATGTGTAGATGAGCCAGGATGGGTCTTGAG	1782
11201	AGGCCAGAGCGGCTGCTTCACTCCGATGTGTAGATGAGCCAGGATGGGTCTTGAG	1260
1783	GAGCCGAGTTCTGAGCTCTGTCTGCACTCGAGCCAGCCACAGCCGCGGATGCACTGCTG	1842
1261	GAGCCGAGTTCTGAGCTCTGTCTGCACTCGAGCCAGCCACAGCCGCGGATGCACTGCTG	1320
1843	GGCAGTTTGTCTGGGAAAACTATATCTCCGAGGCATCTTCCGTATCACGGCTCGGACC	1902
1321	GGCAGTTTGTCTGGGAAAACTATATCTCCGAGGCATCTTCCGTATCACGGCTCGGACC	1380
1903	ACAGCTCTGCAGAACTCATCTCTTTGGAGCAGGCAAGTTGGGTAGAGGTCTGGGG	1962
1381	ACAGCTCTGCAGAACTCATCTCTTTGGAGCAGGCAAGTTGGGTAGAGGTCTGGGG	1440
1963	TTCTCTGAGTCCAGCAGGAGGAATATTTCTACAGATATTTCAAGATGAAAGGCAAGCA	2022
1441	TTCTCTGAGTCCAGCAGGAGGAATATTTCTACAGATATTTCAAGATGAAAGGCAAGCA	1500
2023	ATTAGAGCCTTTAGTTTGGTCAAATCAAAACAAAGAGCTCTGGGCCCTGTGTCTTGTGCC	2082
1501	ATTAGAGCCTTTAGTTTGGTCAAATCAAAACAAAGAGCTCTGGGCCCTGTGTCTTGTGCC	1560
2083	TGGGTGTCCTGGCTGGCTGCACTTGCCTGATGCAAGATGAAGCGGAGGAAAACTC	2142
1561	TGGGTGTCCTGGCTGGCTGCACTTGCCTGATGCAAGATGAAGCGGAGGAAAACTC	1620
2143	ACATGACTTCCAGAGCCACCAACCCCTCTGTCTACATTTACTCTTGGCCAGGCTCTCCAA	2202
1621	ACATGACTTCCAGAGCCACCAACCCCTCTGTCTACATTTACTCTTGGCCAGGCTCTCCAA	1680
2203	GCTCAGGCATTTGGGACCCAGCTCAGAGACCTCTGTCTCTGGCTGCTGAGGCACTCTGG	2262
1681	GCTCAGGCATTTGGGACCCAGCTCAGAGACCTCTGTCTCTGGCTGCTGAGGCACTCTGG	1740
2263	CAAAAAAGACCTTTTTCAGTCCAGATGACCTCAGGAAGCATGGGTTTATGAGTGGGCGCATC	2322

Db 1741 CAAAAAAGACCCCTTTTCAGTCCAGATGACCTCAGGAAGCATGGGTTAGATGGGGCCATC 1800
Qy 2323 ATCTCACCTTCTTGAAGATGGGTATTTCTCAAGACACCCCATCCCTCTGAGCTACAGC 2382
Db 1801 ATCTCACCTTCTTGAAGATGGGTATTTCTCAAGACACCCCATCCCTCTGAGCTACAGC 1860
Qy 2383 TTCAATTCACCTCTGTTTCCAAAGATGTTCTTTGAGCAATGTCCTATGCTTTGGAGGATGAG 2442
Db 1861 TTCAATTCACCTCTGTTTCCAAAGATGTTCTTTGAGCAATGTCCTATGCTTTGGAGGATGAG 1920
Qy 2443 AAGGGAGAGGTAAACATTTCTAATTCATCATATAGATTTTGAAAGACGCTAGAAGCATAT 2502
Db 1921 AAGGGAGAGGTAAACATTTCTAATTCATCATATAGATTTTGAAAGACGCTAGAAGCATAT 1980
Qy 2503 GGAATACATGGCTGTTTGGGCGATCAACACACGTTTCTTATGCGGCTGTTAAGTGAT 2562
Db 1981 GGAATACATGGCTGTTTGGGCGATCAACACACGTTTCTTATGCGGCTGTTAAGTGAT 2040
Qy 2563 GAGGGGAGAGAGATGGAGAAACATTTTCACTGCGGCTGTCTCAGGGGAGGAACCTG 2622
Db 2041 GAGGGGAGAGAGATGGAGAAACATTTTCACTGCGGCTGTCTCAGGGGAGGAACCTG 2100
Qy 2623 ATGCAATGGGTCCCGTCCCTGAGCTGTCTGTCAGCCACATCTCTGAGATCCCTCCAC 2682
Db 2101 ATGCAATGGGTCCCGTCCCTGAGCTGTCTGTCAGCCACATCTCTGAGATCCCTCCAC 2160
Qy 2683 TGCCTGTAGAGACTCGGAACAAACGTTTCTGACACAGTATGCGCCCATTTCCAGAA 2742
Db 2161 TGCCTGTAGAGACTCGGAACAAACGTTTCTGACACAGTATGCGCCCATTTCCAGAA 2220
Qy 2743 ATGGGCATGTGTAGAAAACAGACATGGAGCTCTTAGTGTGACATTTCTGCAATTAATTC 2802
Db 2221 ATGGGCATGTGTAGAAAACAGACATGGAGCTCTTAGTGTGACATTTCTGCAATTAATTC 2280
Qy 2803 AGCCGCCACTGAAGAAGCTTCACTGTATGAGGCGAGCAGACACAGATCAACATGGAGC 2862
Db 2281 AGCCGCCACTGAAGAAGCTTCACTGTATGAGGCGAGCAGCAGACACAGATCAACATGGAGC 2340
Qy 2863 CCCACCATGGTAGTCTGTTCAGGTGGTCCAGTACAGATGCTTATGGCAGATTTCTC 2922
Db 2341 CCCACCATGGTAGTCTGTTCAGGTGGTCCAGTACAGATGCTTATGGCAGATTTCTC 2400
Qy 2923 TTCTCCGTCTCAAGCTCACCAAAACCTGAAGGAGCTGGAACCTAAGTGGAACTCGCTG 2982
Db 2401 TTCTCCGTCTCAAGCTCACCAAAACCTGAAGGAGCTGGAACCTAAGTGGAACTCGCTG 2460
Qy 2983 AGCCACTCTGCAAGTGAAGAGTCTTTGTAAAGCCCTGAGACGCCCTCGCTCTCTGGAG 3042
Db 2461 AGCCACTCTGCAAGTGAAGAGTCTTTGTAAAGCCCTGAGACGCCCTCGCTCTCTGGAG 2520
Qy 3043 ACCCTCGGTTGGCTGGCTGTGGCTTCACAGCTGAGGACTGCAAGGACCTTGGCTTTGGG 3102
Db 2521 ACCCTCGGTTGGCTGGCTGTGGCTTCACAGCTGAGGACTGCAAGGACCTTGGCTTTGGG 2580
Qy 3103 CTGAGAGCCAAACAGACCTTGACCGAGCTGGACCTGAGCTTCAATGTGCTCACGGATGCT 3162
Db 2581 CTGAGAGCCAAACAGACCTTGACCGAGCTGGACCTGAGCTTCAATGTGCTCACGGATGCT 2640
Qy 3163 GGAGCCAAACACTTTTGGCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCGACTGCGAG 3222
Db 2641 GGAGCCAAACACTTTTGGCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCGACTGCGAG 2700
Qy 3223 CTGGTCAGTGTGGCTCTCAGCTGTGATCTGCTGCGCAGGACTGCGCTCTGTGCTTAGTGCC 3282
Db 2701 CTGGTCAGTGTGGCTCTCAGCTGTGATCTGCTGCGCAGGACTGCGCTCTGTGCTTAGTGCC 2760
Qy 3283 AGCCCCAGCCTGAAGGAGCTAGACCTTGACGAGAACCAACTGGATGACGCTTGGCGTGGGA 3342
Db 2761 AGCCCCAGCCTGAAGGAGCTAGACCTTGACGAGAACCAACTGGATGACGCTTGGCGTGGGA 2820
Qy 3343 CTGCTCTGTGAGGGGCTCAGGATCTCTGCTGCAAACTCATACGCTTGGGGCTGGACCGAG 3402
Db 2821 CTGCTCTGTGAGGGGCTCAGGATCTCTGCTGCAAACTCATACGCTTGGGGCTGGACCGAG 2880

Qy 3403 ACAACTCTGAGTGTAGATGAGGAGGAACTGAGGGCCCTGGAGCAGGAGAACTCTCAG 3462
Db 2881 ACAACTCTGAGTGTAGATGAGGAGGAACTGAGGGCCCTGGAGCAGGAGAACTCTCAG 2940
Qy 3463 CTGCTCATCTTTCAGCAGACGAAACAAAGTGTGTAGTACCCCTACTGTAGGGGCTGGATACG 3522
Db 2941 CTGCTCATCTTTCAGCAGACGAAACAAAGTGTGTAGTACCCCTACTGTAGGGGCTGGATACG 3000
Qy 3523 GGAGAGATGAGTAATAGACATCTCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGGCG 3582
Db 3001 GGAGAGATGAGTAATAGACATCTCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGGCG 3060
Qy 3583 GCTTCCCATGTTGCTCAGGCTAACTCTCAAACTCTCTGGACGTGAGCAAGATCTTCCCAATT 3642
Db 3061 GCTTCCCATGTTGCTCAGGCTAACTCTCAAACTCTCTGGACGTGAGCAAGATCTTCCCAATT 3120
Qy 3643 GCTGAGATTGAGAGGAAAGCTCCCTCAGAGGTAGTACCGGTGGAACTCTTGTGCGTGCT 3702
Db 3121 GCTGAGATTGAGAGGAAAGCTCCCTCAGAGGTAGTACCGGTGGAACTCTTGTGCGTGCT 3180
Qy 3703 TCTCTGCTCTCAAGGGGACCTGCTATACGAGCTTTTGGGGACTGACGATGACTTCTGG 3762
Db 3181 TCTCTGCTCTCAAGGGGACCTGCTATACGAGCTTTTGGGGACTGACGATGACTTCTGG 3240
Qy 3763 GGCCCGACGGGCTGTGGCTACTGAGGTAGTGTGACAAAGAAAGAACTTGTACCGAGTT 3822
Db 3241 GGCCCGACGGGCTGTGGCTACTGAGGTAGTGTGACAAAGAAAGAACTTGTACCGAGTT 3300
Qy 3823 CACTTCCCTGTAGCTGGCTCTCTACCGTGGCCCAACACGGGTCTCTGTGTTGTGATGAGA 3882
Db 3301 CACTTCCCTGTAGCTGGCTCTCTACCGTGGCCCAACACGGGTCTCTGTGTTGTGATGAGA 3360
Qy 3883 GAAGCGGTGACGTTGAGATTGAATTTCTGTGTGGGACCAAGTCTCTGGGTGAGATCAAC 3942
Db 3361 GAAGCGGTGACGTTGAGATTGAATTTCTGTGTGGGACCAAGTCTCTGGGTGAGATCAAC 3420
Qy 3943 CCACAGCAGCTGATGGTGGCAGGGCTCTGCTGGACATCAAGGCTGAGCCTGGAGCT 4002
Db 3421 CCACAGCAGCTGATGGTGGCAGGGCTCTGCTGGACATCAAGGCTGAGCCTGGAGCT 3480
Qy 4003 GTGGAAGCTGTGCACCTCCCTCACTTTGTGGCTCTTCCAAGGGGGCCATGTGAGACATCC 4062
Db 3481 GTGGAAGCTGTGCACCTCCCTCACTTTGTGGCTCTTCCAAGGGGGCCATGTGAGACATCC 3540
Qy 4063 CTGTTCCAAATGGCCCACTTTAAAGAGGAGGGGATGCTCTCTGGAGAACCCAGGGGTG 4122
Db 3541 CTGTTCCAAATGGCCCACTTTAAAGAGGAGGGGATGCTCTCTGGAGAACCCAGGGGTG 3600
Qy 4123 GAGCTGCATCATAGTTTCTGGAAACCCAGCTTCTCCCTCTGGGAGTCTCTCTGAAA 4182
Db 3601 GAGCTGCATCATAGTTTCTGGAAACCCAGCTTCTCCCTCTGGGAGTCTCTCTGAAA 3660
Qy 4183 ATGATCCATAATGCGCTCGCTTCATTTCCCGTCACTCTGTGTGTGTTTACACCGC 4242
Db 3661 ATGATCCATAATGCGCTCGCTTCATTTCCCGTCACTCTGTGTGTGTTTACACCGC 3720
Qy 4243 GTCCATCTGAGGAGTCACTTCACTCTACCTGATCCCAAGTACTGCTTCCATTGG 4302
Db 3721 GTCCATCTGAGGAGTCACTTCACTCTACCTGATCCCAAGTACTGCTTCCATTGG 3780
Qy 4303 AAGGAACCTGAGGCTGCTGCTATGAGGACCCCTGAGAGACCAAGCTGTTCTCGAGTTCTAC 4362
Db 3781 AAGGAACCTGAGGCTGCTGCTATGAGGACCCCTGAGAGACCAAGCTGTTCTCGAGTTCTAC 3840
Qy 4363 GTTGGCCACTTGGGATCAGGATCAGGCTGCAAGTGAAGAAAGAAAGATGAGACTCTG 4422
Db 3841 GTTGGCCACTTGGGATCAGGATCAGGCTGCAAGTGAAGAAAGAAAGATGAGACTCTG 3900
Qy 4423 GTTGGGAGGCTTGGTGAACCCAGGAGATCTATGCTGTGCAACTATCTGATCTCTCCA 4482
Db 3901 GTTGGGAGGCTTGGTGAACCCAGGAGATCTATGCTGTGCAACTATCTGATCTCTCTCCA 3960

```
QY 4483 GCGCGATAGCCGTACTTCACTCTGGATGCCCGCAGTTGCTGCACTTTGTGACCAAG 4542
Db |||||
QY 3961 GCGCGATAGCCGTACTTCACTCTGGATGCCCGCAGTTGCTGCACTTTGTGACCAAG 4020
Db |||||
QY 4543 TATCGAGACAGCTGATAGCCCGAGTGCATCGGTGGAGTTGCTTGGACAACTGCAT 4602
Db |||||
QY 4021 TATCGAGACAGCTGATAGCCCGAGTGCATCGGTGGAGTTGCTTGGACAACTGCAT 4080
QY 4603 GGAACAGTCTGAGCCAGGAGCAGTACGAGAGGGTCTGGCTCAGAAACAGAGGCCAGC 4662
Db |||||
QY 4081 GGACAGTCTGAGCCAGGAGCAGTACGAGAGGGTCTGGCTGAGAAACAGAGGCCAGC 4140
QY 4663 CAGATCGGAAAGCTGTTCAGCTTGGAGCCAGTCTCTGGGACCGGAAGTGCAGAACTGCACTC 4722
Db |||||
QY 4141 CAGATCGGAAAGCTGTTCAGCTTGGAGCCAGTCTCTGGGACCGGAAGTGCAGAACTGCACTC 4200
QY 4723 TACCAAGCCCTGAAGAGAGCCATCTCCTACCTCATTTATGAACTCTGGGAGAGGGCAGC 4782
Db |||||
QY 4201 TACCAAGCCCTGAAGAGAGCCATCTCCTACCTCATTTATGAACTCTGGGAGAGGGCAGC 4260
QY 4783 AAAAAAGGAGTCTCTGCACTCAGCAGC 4809
Db |||||
QY 4261 AAAAAAGGAGTCTCTGCACTCAGCAGC 4287

RESULT 7
US-09-931-071-5
; Sequence 5, Application US/09931071
; Patent No. US20020128219A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; APPLICANT: Alnemri, Emad S.
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-335001
; CURRENT APPLICATION NUMBER: US/09/931,071
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/428,252
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4287
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-931-071-5

Query Match 78.7%; Score 4287; DB 10; Length 4287;
Best Local Similarity 100.0%; Pred. No.: 0;
Matches 4287; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 523 ATGGCTGGCGGAGCCTGGGCGCCCTGGCTGTACTTGGAGTTCCTGAGAGGAGGAG 582
Db 1 ATGGCTGGCGGAGCCTGGGCGCCCTGGCTGTACTTGGAGTTCCTGAGAGGAGGAG 60
QY 583 CTGAAGAGTTCAGCTTCTGCTCGCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG 642
Db 61 CTGAAGAGTTCAGCTTCTGCTCGCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG 120
QY 643 ACACCCGCTCAGCCAGAGAAACGAGTGGCATGGAGTGGCTCTGCTAGCTGGTCTAG 702
Db 121 ACACCCGCTCAGCCAGAGAAACGAGTGGCATGGAGTGGCTCTGCTAGCTGGTCTAG 180
QY 703 TATGGGAGCAGCGGGCTGGGACTTACCTCCTCATACCTGGGAGCAGATGGGCTGAG 762
Db 181 TATGGGAGCAGCGGGCTGGGACTTACCTCCTCATACCTGGGAGCAGATGGGCTGAG 240
QY 763 TCACTGTGCGCCCAAGCCAGGAGGGGAGGCGCACTCTCCCTCATTTCCCTACAGCCCA 822
Db 241 TCACTGTGCGCCCAAGCCAGGAGGGGAGGCGCACTCTCCCTCATTTCCCTACAGCCCA 300
QY 823 AGTGAACCCCACTGGGGTCTCCGAGCAACCCACCTCCACCGCAGTCTAATGCCCTGG 882
Db |||||
```

```
Db 301 AGTGAACCCCACTGGGGTCTCCGAGCAACCCACCTCCACCGCAGTGTCTAATGCCCTGG 360
QY 883 ATCCATGAATTCGCGGGGTGCACCCAGGGCTCAGAGAGAGGGTTTGTGAGACAGCTG 942
Db |||||
QY 361 ATCCATGAATTCGCGGGGTGCACCCAGGGCTCAGAGAGAGGGTTTGTGAGACAGCTG 420
QY 943 CTGACACATCTGACGCGCGCTGGAGAGAAATCTCTGCTCCTCTCTACCAAGCTCTT 1002
Db 421 CTGACACATCTGACGCGCGCTGGAGAGAAATCTCTGCTCCTCTCTCTACCAAGCTCTT 480
QY 1003 CCAAGTCTCCGAGACCATGAGTCTCCAAAGCCAGAGTCAACCAAGCCGCCCCACATCCACA 1062
Db 481 CCAAGTCTCCGAGACCATGAGTCTCCAAAGCCAGAGTCAACCAAGCCGCCCCACATCCACA 540
QY 1063 GCAGTGTGGGAGCTGGGGATCCCCACCTCAGCCCCAGCTAGCACCCAGAGAGCAGAG 1122
Db 541 GCAGTGTGGGAGCTGGGGATCCCCACCTCAGCCCCAGCTAGCACCCAGAGAGCAGAG 600
QY 1123 GCTCTGGGACCCAAATGSCCTCTGGATGAAACGTCAGGAATTTACTACAGAAATCAGA 1182
Db 601 GCTCTGGGACCCAAATGSCCTCTGGATGAAACGTCAGGAATTTACTACAGAAATCAGA 660
QY 1183 GAAAGAGAGAGAGAAATCAGAAAGGAGGCGGCCCATGGGAGCGGTGGTGAAGACG 1242
Db 661 GAAAGAGAGAGAGAAATCAGAAAGGAGGCGGCCCATGGGAGCGGTGGTGAAGACG 720
QY 1243 CCCCCACAGGGCCACACCCAGCTACAGCCCCACACACCCATGGGAGCCTTCTGTGAGA 1302
Db 721 CCCCCACAGGGCCACACCCAGCTACAGCCCCACACACCCATGGGAGCCTTCTGTGAGA 780
QY 1303 GAGAGCCTCTGTTCCACATGGCCCTGGAAATGAGGATTTTAAACCAAAATTCACACAG 1362
Db 781 GAGAGCCTCTGTTCCACATGGCCCTGGAAATGAGGATTTTAAACCAAAATTCACACAG 840
QY 1363 CTGCTACTTCTACAAAGACCTCAACCCAGAGCCAGATCCCTGCTCAAGAGAGCTGG 1422
Db 841 CTGCTACTTCTACAAAGACCTCAACCCAGAGCCAGATCCCTGCTCAAGAGAGCTGG 900
QY 1423 CCGTATATTGTGGAGGAGATCGAGGACATTTAATTGAGATCAGAGACTTATTTGGCCCA 1482
Db 901 CCGTATATTGTGGAGGAGATCGAGGACATTTAATTGAGATCAGAGACTTATTTGGCCCA 960
QY 1483 GGCCTGGATACCCAAAGAACCTCGCATATCTATGACAGGGGCTGCTGGAAATTTGGGAG 1542
Db 961 GGCCTGGATACCCAAAGAACCTCGCATATCTATGACAGGGGCTGCTGGAAATTTGGGAG 1020
QY 1543 TCAACACTGGCCAGGCTGAAGAGCCTGGGGAGAGGCCAGCTCTATGGGACCGC 1602
Db 1021 TCAACACTGGCCAGGCTGAAGAGCCTGGGGAGAGGCCAGCTCTATGGGACCGC 1080
QY 1603 TTCCAGCATGCTTCTTACTCTAGCTGAGAGAGCTGGGCCAGTCCAAAGTGTGAGTCTC 1662
Db 1081 TTCCAGCATGCTTCTTACTCTAGCTGAGAGAGCTGGGCCAGTCCAAAGTGTGAGTCTC 1140
QY 1663 GCTGAGCTCATCGGAAAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCTGTCT 1722
Db 1141 GCTGAGCTCATCGGAAAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCTGTCT 1200
QY 1723 AGGCCAGAGCGGCTCTCTTCTATCTCTGATGGTGTAGATGAGCCAGGATGGGTCTTGAG 1782
Db 1201 AGGCCAGAGCGGCTCTCTTCTATCTCTGATGGTGTAGATGAGCCAGGATGGGTCTTGAG 1260
QY 1783 GAGCCAGTCTGAGCTCTGCTGCACTGGAGCCAGCCACAGCCGGCGGATGCATCTGCTG 1842
Db 1261 GAGCCAGTCTGAGCTCTGCTGCACTGGAGCCAGCCACAGCCGGCGGATGCATCTGCTG 1320
QY 1843 GGCAGTTTGTGGGAAAAATATATCTTCCGAGGATCTCTTCTGATCAGCGCTCGGACC 1902
Db 1321 GGCAGTTTGTGGGAAAAATATATCTTCCGAGGATCTCTTCTGATCAGCGCTCGGACC 1380
QY 1903 ACAGCTCTGCAAACTCATTTCTTTTGGAGAGGCAAGTGGGTGAGGTCTCTGGG 1962
Db 1381 ACAGCTCTGCAAACTCATTTCTTTTGGAGAGGCAAGTGGGTGAGGTCTCTGGG 1440
```

Qy	1963	TTCTCTGAGTCCAGCAGGAAAGGAATATTTCTACAGATATTTTCACAGATGAAAGCAAGCA	2022
Db	1441	TTCTCTGAGTCCAGCAGGAAAGGAATAATTTCTACAGATATTTTCACAGATGAAAGCAAGCA	1500
Qy	2023	ATTAGAGCCTTTAGGTTGGTCAAAATCAAAACAAAGAGCTCTGGGCCCTGTGTCTTTGTGCC	2082
Db	1501	ATTAGAGCCTTTAGGTTGGTCAAAATCAAAACAAAGAGCTCTGGGCCCTGTGTCTTTGTGCC	1560
Qy	2083	TGGGTGCTGGCTGGCCCTGCACCTTGCTGATGCAGCAGATGAAGCGGAAGGAAAAATC	2142
Db	1561	TGGGTGCTGGCTGGCCCTGCACCTTGCTGATGCAGCAGATGAAGCGGAAGGAAAAATC	1620
Qy	2143	ACACTGACTTCCAAGACCAACACAAACCTCTGTCTACTATTAACCTTGCCCAAGGCTCTCCAA	2202
Db	1621	ACACTGACTTCCAAGACCAACACAAACCTCTGTCTACTATTAACCTTGCCCAAGGCTCTCCAA	1680
Qy	2203	GCTCAGCCATTGGGACCCCAAGCTCAGAGACCTCTGTCTCTGTGCTGTGAGGCGCATCTGG	2262
Db	1681	GCTCAGCCATTGGGACCCCAAGCTCAGAGACCTCTGTCTCTGTGCTGTGAGGCGCATCTGG	1740
Qy	2263	CAAAAAAGACCCCTTTTCAGTCCAGATGACCTCAGGAAGCATGGGTTAGATGGGCCATC	2322
Db	1741	CAAAAAAGACCCCTTTTCAGTCCAGATGACCTCAGGAAGCATGGGTTAGATGGGCCATC	1800
Qy	2323	ATCTCCACCTTTCTTGAAGATGGGTATTTCTCAAGAGCACCCATCCCTCTTGAGCTACAGC	2382
Db	1801	ATCTCCACCTTTCTTGAAGATGGGTATTTCTCAAGAGCACCCATCCCTCTTGAGCTACAGC	1860
Qy	2383	TTCAATTCACCTCTGTTTTCCAAGAGTTCCTTTGCACAAATGTCCTATGTCTTTGGAGTAG	2442
Db	1861	TTCAATTCACCTCTGTTTTCCAAGAGTTCCTTTGCACAAATGTCCTATGTCTTTGGAGTAG	1920
Qy	2443	AAGGGAGAGGTAAACATTTCTAATTGCATCATAGATTTGAAAAGAGCGCTAGAACATAT	2502
Db	1921	AAGGGAGAGGTAAACATTTCTAATTGCATCATAGATTTGAAAAGAGCGCTAGAACATAT	1980
Qy	2503	GGAATACATGSCCTGTTTGGGGCATCAACACACGTTTTCTATTGGGCCCTGTTTAAGTGAT	2562
Db	1981	GGAATACATGSCCTGTTTGGGGCATCAACACACGTTTTCTATTGGGCCCTGTTTAAGTGAT	2040
Qy	2563	GAGGGGAGAGAGAGATGGAGAAATCTTTCACTGCCGGCTGTCTCAGGGGAGGAAACCTG	2622
Db	2041	GAGGGGAGAGAGAGATGGAGAAATCTTTCACTGCCGGCTGTCTCAGGGGAGGAAACCTG	2100
Qy	2623	ATGCAGTGGGTCCGTCCTGTGAGCTGCTGTGCAGCCACACTCTCTGAGAGTCCCTCCAC	2682
Db	2101	ATGCAGTGGGTCCGTCCTGTGAGCTGCTGTGCAGCCACACTCTCTCTGAGAGTCCCTCCAC	2160
Qy	2683	TGCTTGTACGAGACTCGGAACAAAACGTTCTCTGCACACAAGTGATGGGCCCATTTTCGAGAA	2742
Db	2161	TGCTTGTACGAGACTCGGAACAAAACGTTCTCTGCACACAAGTGATGGGCCCATTTTCGAGAA	2220
Qy	2743	ATGGGCATGTGTAGAAAACAGACATGGAGCTCTTAGTGTGCATTTCTGCATTTAAATTC	2802
Db	2221	ATGGGCATGTGTAGAAAACAGACATGGAGCTCTTAGTGTGCATTTCTGCATTTAAATTC	2280
Qy	2803	AGCCGCCAGTGGAAGAGCTTCACTAGCTGATTAAGGCGAGGCGAGCACAGATCAACATGGAGC	2862
Db	2281	AGCCGCCAGTGGAAGAGCTTCACTAGCTGATTAAGGCGAGGCGAGCACAGATCAACATGGAGC	2340
Qy	2863	CCNACCATGTAGTCTGTTTCAGGTGGGTCCAGTCAAGATGCCTATTGGCAGATTCCTC	2922
Db	2341	CCNACCATGTAGTCTGTTTCAGGTGGGTCCAGTCAAGATGCCTATTGGCAGATTCCTC	2400
Qy	2923	TTCTCCGCTCTCAAGGTCAACAGAAACCTGAAGGAGCTGCACTTAAGTGGAAACTCGCTG	2982
Db	2401	TTCTCCGCTCTCAAGGTCAACAGAAACCTGAAGGAGCTGCACTTAAGTGGAAACTCGCTG	2460
Qy	2983	AGCCACTCTCAGTGAAGAGTCTTTGTAAGACCTTGAGACGCCCTCGCTCCCTCTCTGGAG	3042
Db	2461	AGCCACTCTCAGTGAAGAGTCTTTGTAAGACCTTGAGACGCCCTCGCTCCCTCTCTGGAG	2520

QY	3043	ACCTCGGGTTGCTTGGCTGTGGCTCACAAGCTTGAGGACTTGCAAGNACCTTGCTTTGGG	3107
DB	2521	ACCCTGCGGTTGGCTGTGGCTCACAAGCTTGAGGACTTGCAAGNACCTTGCTTTGGG	2580
QY	3103	CTGAGAGCCAAACAGACCCCTGACCGAGCTGGAACCTTGAGCTTCAAATGTGCTCACGGATGCT	3162
DB	2581	CTGAGAGCCAAACAGACCCCTGACCGAGCTGGAACCTTGAGCTTCAAATGTGCTCACGGATGCT	2640
QY	3163	GGAGCCAAAACACCTTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCGACTGCAG	3222
DB	2641	GGAGCCAAAACACCTTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCGACTGCAG	2700
QY	3223	CTGCTCAGCTGTGGCCTCA CGTCTGACTGTCGCCAGSACCTGGCCTCTGTGCTTAGTGCC	3282
DB	2701	CTGCTCAGCTGTGGCCTCA CGTCTGACTGTCGCCAGSACCTGGCCTCTGTGCTTAGTGCC	2760
QY	3283	AGCCCCAGCCTGAAGGAGCTTAGACCTGCGACGAGAACAACTCGATGACGCTTGCGCTGCGA	3342
DB	2761	AGCCCCAGCCTGAAGGAGCTTAGACCTGCGACGAGAACAACTCGATGACGCTTGCGCTGCGA	2820
QY	3343	CTGCTCTGTGAGGGGCTCAGGCATCTCGCTCGCCGAAACTCATAGCCCTGGSGGCTGGACCAG	3402
DB	2821	CTGCTCTGTGAGGGGCTCAGGCATCTCGCTCGCCGAAACTCATAGCCCTGGSGGCTGGACCAG	2880
QY	3403	ACAACTCTGAGTGATGATGAGGAGGAGGAACTGAGGGCCCTGAGCAGGAGAGAACTCTCAG	3462
DB	2881	ACAACTCTGAGTGATGATGAGGAGGAGGAACTGAGGGCCCTGAGCAGGAGAGAACTCTCAG	2940
QY	3463	CTGCTCATCTTCAGCAGACGGAAACCAAGTGTGATGACCCCCTACTGAGGGCCTGGATACG	3522
DB	2941	CTGCTCATCTTCAGCAGACGGAAACCAAGTGTGATGACCCCCTACTGAGGGCCTGGATACG	3000
QY	3523	GGAGAGATGAGTAATAGCACATCTCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGGCG	3582
DB	3001	GGAGAGATGAGTAATAGCACATCTCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGGCG	3060
QY	3583	GCTTCCCATGTGTCTCAGGCTAATCTCAAACTCCTGGACGTGAGCAAGATCTTCCCAATT	3642
DB	3061	GCTTCCCATGTGTGTCTCAGGCTAATCTCAAACTCCTGGACGTGAGCAAGATCTTCCCAATT	3120
QY	3643	GCTGAGATTGCAAGGAAAGCTCCCAAGAGGTAGTACCCGTTGAGAACCTCTTGTGCGTGCCT	3702
DB	3121	GCTGAGATTGCAAGGAAAGCTCCCAAGAGGTAGTACCCGTTGAGAACCTCTTGTGCGTGCCT	3180
QY	3703	TCTCCTCCCTCTCAAGGGGACCTGCTACCAAGCCTTTGGGGACTGACGATGACTTCTTGG	3762
DB	3181	TCTCCTCCCTCTCAAGGGGACCTGCTACCAAGCCTTTGGGGACTGACGATGACTTCTTGG	3240
QY	3763	GGCCCCACGGGGCCTGTGGCTACTGAGGTAGTTTGACAAAAAGAACTTTGTACCGAGTT	3822
DB	3241	GGCCCCACGGGGCCTGTGGCTACTGAGGTAGTTTGACAAAAAGAACTTTGTACCGAGTT	3300
QY	3823	CAC TTCCCTGTAGCTTGCTCCTACCGCTGCCCCAAC CACGGGTCTCTGCTTTGTGATGAGA	3882
DB	3301	CAC TTCCCTGTAGCTTGCTCCTACCGCTGCCCCAAC CACGGGTCTCTGCTTTGTGATGAGA	3360
QY	3883	GNAGCGGTGACCGTTGAGATTGAATCTGTGTGTGGGACCGACTCTCTGGGTGAGATCAAC	3942
DB	3361	GNAGCGGTGACCGTTGAGATTGAATCTGTGTGTGGGACCGACTCTCTGGGTGAGATCAAC	3420
QY	3943	CCACAGCACAGCTGGATGTGGGAGGGCCTCTGCTGGACATCAAGGCTGAGCCTGGAGCT	4002
DB	3421	CCACAGCACAGCTGGATGTGGGAGGGCCTCTGCTGGACATCAAGGCTGAGCCTGGAGCT	3480
QY	4003	GTGGAAGCTGTGACCTCTCCATCTTTGTGGCTCTCCAAGGGGCCNTGTGGACACATCC	4062
DB	3481	GTGGAAGCTGTGACCTCTCCATCTTTGTGGCTCTCCAAGGGGCCNTGTGGACACATCC	3540
QY	4063	CTGTTCCAAATGGCCCACTTTAAAGAGGAGGGGATGCTCTCTGGAGAGCCAGCCAGGGTG	4122
DB	3541	CTGTTCCAAATGGCCCACTTTAAAGAGGAGGGGATGCTCTCTGGAGAGCCAGCCAGGGTG	3600
QY	4123	GAGCTGCATCACATPAGTCTTGGAAAAACCCAGGCTTCTCCCTCTGGAGGTCTCTCTGAAA	4182

Db 3601 |||||GAGCTGCATCACATAGTTCTGGAACCCAGCTTCTCCCTTGGAGTCTCTCTGAAA 3660
QY 4183 ATGATCCATAATGCGCTGCGCTTCAATTCGGCTCACCTCTGTGTGTGCTTTACACCGC 4242
Db 3661 ATGATCCATAATGCGCTGCGCTTCAATTCGGCTCACCTCTGTGTGTGCTTTACACCGC 3720
QY 4243 GTCCATCTCTGAGGAAGTCACTTCCACCTCTACCTGATCCCAAGTGAAGTCTTCCATTCGG 4302
Db 3721 GTCCATCTCTGAGGAAGTCACTTCCACCTCTACCTGATCCCAAGTGAAGTCTTCCATTCGG 3780
QY 4303 AAGGAAGTGGAGCTCTGCTATCGAAGCCCTGGAAGACAGCTGTTCTCGGAGTCTTAC 4362
Db 3781 AAGGAAGTGGAGCTCTGCTATCGAAGCCCTGGAAGACAGCTGTTCTCGGAGTCTTAC 3840
QY 4363 GTTGGCCACTTGGGATCAGGGATCAGGCTGCAAGTGAAGACAGAAAGTGAAGTCTG 4422
Db 3841 GTTGGCCACTTGGGATCAGGGATCAGGCTGCAAGTGAAGACAGAAAGTGAAGTCTG 3900
QY 4423 GTTGGGAGGCTTGGTGAACCCAGAGATCTCATGCTTCACTCTGATCCCTCCA 4482
Db 3901 GTTGGGAGGCTTGGTGAACCCAGAGATCTCATGCTTCACTCTGATCCCTCCA 3960
QY 4483 GCCGATAGCCGATACCTTCACTCTGATGCCCGCAGTTGTCGACTTTTGTGACACG 4542
Db 3961 GCCGATAGCCGATACCTTCACTCTGATGCCCGCAGTTGTCGACTTTTGTGACACG 4020
QY 4543 TATCGAGACGCTGATAGCCGAGTGACATCGGTGAGGTTGCTTGGACAACTGCAT 4602
Db 4021 TATCGAGACGCTGATAGCCGAGTGACATCGGTGAGGTTGCTTGGACAACTGCAT 4080
QY 4603 GGACAGGCTCTGACCCAGGACGATACGAGAGGCTCTGCTCAGAAACAGAGGCCCGC 4662
Db 4081 GGACAGGCTCTGACCCAGGACGATACGAGAGGCTCTGCTCAGAAACAGAGGCCCGC 4140
QY 4663 CAGATCGGAAGCTGTTTCAAGTTGAGCGAGTCTCTGGGACGGAAAGTGCAAGTGA 4722
Db 4141 CAGATCGGAAGCTGTTTCAAGTTGAGCGAGTCTCTGGGACGGAAAGTGCAAGTGA 4200
QY 4723 TACCAAGCCCTGAAGGAGACCCATCTCACCTATTATGGAATCTGCGAGAGGCGCAG 4782
Db 4201 TACCAAGCCCTGAAGGAGACCCATCTCACCTATTATGGAATCTGCGAGAGGCGCAG 4260
QY 4783 AAAAAAGGACTCTGCGCACTCAGCAGC 4809
Db 4261 AAAAAAGGACTCTGCGCACTCAGCAGC 4287

RESULT 8
US-09-388-221-1
; Sequence 1, Application US/09388221A
; Publication No. US20020192643A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. US20020192643A1el Card Proteins Involved in Cell Death Regul
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388, 221A
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4422
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(4422)
US-09-388-221-1

Query Match 76.2%; Score 4148; DB 10; Length 4422;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 4290; Conservative 0; Mismatches 0; Indels 132; Gaps 1;

QY 523 ATGCTGCGGAGCTGCGGCGCTGCTGCTGTTACTTGGAGTTCCTCGAAGAGAGGAG 582
Db 1 ATGCTGCGGAGCTGCGGCGCTGCTGCTGTTACTTGGAGTTCCTCGAAGAGAGGAG 60
QY 583 CTGAGGAGTTCAGCTTCTGCTCGCCAAATAAGCGCACTCCAGGAGCTCTTCGGGTGAG 642
Db 61 CTGAGGAGTTCAGCTTCTGCTCGCCAAATAAGCGCACTCCAGGAGCTCTTCGGGTGAG 120
QY 643 ACACCCGCTCAGCCAGAGAAAGCAGAGTGGCCTGAGAGTGGCCTCGTACTCGTGGCTCAG 702
Db 121 ACACCCGCTCAGCCAGAGAAAGCAGAGTGGCCTGAGAGTGGCCTCGTACTCGTGGCTCAG 180
QY 703 TATGGGAGAGCGGGCTTGGGACCTAGCCCTCCATCTGCGGAGCAGATGGGCTGAGG 762
Db 181 TATGGGAGAGCGGGCTTGGGACCTAGCCCTCCATCTGCGGAGCAGATGGGCTGAGG 240
QY 763 TCACCTGCGGCCCAAGCCAGAGAGGGGCGAGGCCACTCTCCCTCATTCCTCCCTACAGCCCA 822
Db 241 TCACCTGCGGCCCAAGCCAGAGAGGGGCGAGGCCACTCTCCCTCATTCCTCCCTACAGCCCA 300
QY 823 AGTGAACCCCACTTGGGCTCTCCAGCCAAACCCACTCCACCGCAGTGTATGTCCTGG 882
Db 301 AGTGAACCCCACTTGGGCTCTCCAGCCAAACCCACTCCACCGCAGTGTATGTCCTGG 360
QY 883 ATCCATGAATTCGCGCGGGTGCACCCAGGCTCAGAGAGAGGGTTTGTGAGACAGCTG 942
Db 361 ATCCATGAATTCGCGCGGGTGCACCCAGGCTCAGAGAGAGGGTTTGTGAGACAGCTG 420
QY 943 CTGACACATCTGGAGCGGCTGGAGAGAAATCTCTGCTCCTCCTCTACCAAGCTCTT 1002
Db 421 CTGACACATCTGGAGCGGCTGGAGAGAAATCTCTGCTCCTCCTCTACCAAGCTCTT 480
QY 1003 CCAAGCTCCCAAGACCATGAGTCTCCAAGCCAGGAGTCAACCAAGCCGCCCATCTCCACA 1062
Db 481 CCAAGCTCCCAAGACCATGAGTCTCCAAGCCAGGAGTCAACCAAGCCGCCCATCTCCACA 540
QY 1063 GCAGTCTGGGAGCTGGGGATCCCACTCAGCCAGCTAGACACCCAGAGAGGAG 1122
Db 541 GCAGTCTGGGAGCTGGGGATCCCACTCAGCCAGCTAGACACCCAGAGAGGAG 600
QY 1123 GCTCTGGGACCCCAATGSCCTCTGGATGAACGTCAGGAAATTTACTACACAGAAATCAGA 1182
Db 601 GCTCTGGGACCCCAATGSCCTCTGGATGAACGTCAGGAAATTTACTACACAGAAATCAGA 660
QY 1183 GAAAGAGAGAGAGAAATCAGAGAAAGCAGGCCCTCATGGGCGAGCGTGTGAGAAACG 1242
Db 661 GAAAGAGAGAGAGAAATCAGAGAAAGCAGGCCCTCATGGGCGAGCGTGTGAGAAACG 720
QY 1243 CCCCCAGAGGCGCACACAGGCTTACAGCCCAACCAACCCATGGGAGCCTTCTGTGAGA 1302
Db 721 CCCCCAGAGGCGCACACAGGCTTACAGCCCAACCAACCCATGGGAGCCTTCTGTGAGA 780
QY 1303 GAGAGCCTCTGTTCCACATGGCCCTGGAAATGAGGATTTTAAACCAAAATTCACACAG 1362
Db 781 GAGAGCCTCTGTTCCACATGGCCCTGGAAATGAGGATTTTAAACCAAAATTCACACAG 840
QY 1363 CTGCTACTTTTACAAAGACCTCACCCCAAGAGTCCCTGCTGCTCAAGAGAAAGCTGG 1422
Db 841 CTGCTACTTTTACAAAGACCTCACCCCAAGAGTCCCTGCTGCTCAAGAGAAAGCTGG 900
QY 1423 CCTGATTTGTGGAGAGAAATCAGAGCAATTTAAATTGAGATCAGAGCTTATTTGGCCCA 1482
Db 901 CCTGATTTGTGGAGAGAAATCAGAGCAATTTAAATTGAGATCAGAGCTTATTTGGCCCA 960
QY 1483 GGCCTGGATACCCAAAGACCTCGATAGTATCTAGCGAGGGGCTGCTGGAAATGGGAAG 1542
Db 961 GGCCTGGATACCCAAAGACCTCGATAGTATCTAGCGAGGGGCTGCTGGAAATGGGAAG 1020
QY 1543 TCAACACTGGCCAGGCGAGTGAAGAGCCTCGGGGAGAGGCGAGCTGTATGGGACCGC 1602
Db 1021 TCAACACTGGCCAGGCGAGTGAAGAGCCTCGGGGAGAGGCGAGCTGTATGGGACCGC 1080
QY 1603 TTCCAGCATGTCTTCTACTTTCAGCTGCGAGAGAGCTGGGCCAGTCCCAAGGTTGTGAGTCTC 1662

1081	Db	 TTCCAGCATGCTCTTACATTACAGTGCAGAGCTGGCCCAAGTGGTGAAGTCTC	1140
1663	Qy	GCTGAGCTCATCGGAAAGATGGGACAGCACTCCGGCTCCCAATTAGACAGATCCTGTCT	1722
1141	Db	GCTGAGCTCATCGGAAAGATGGGACAGCACTCCGGCTCCCAATTAGACAGATCCTGTCT	1200
1723	Qy	AGCCAGAGCGGCTGCTCTTCACTCCTCGATGGTGTAGATGAGCCAGGATGGGTTTGCAG	1782
1201	Db	AGCCAGAGCGGCTGCTCTTCACTCCTCGATGGTGTAGATGAGCCAGGATGGGTTTGCAG	1260
1783	Qy	GAGCCAGTTCTGAGCTCTGTCTGCACTGGAGCCAGCCACAGCCGGCGGATGCACTGCTG	1842
1261	Db	GAGCCAGTTCTGAGCTCTGTCTGCACTGGAGCCAGCCACAGCCGGCGGATGCACTGCTG	1320
1843	Qy	GGCAGTTGCTGGGAAACTATCTTCCGAGGCATCTTCTTCTGATCAGCGCTCGGACC	1902
1321	Db	GGCAGTTGCTGGGAAACTATCTTCCGAGGCATCTTCTTCTGATCAGCGCTCGGACC	1380
1903	Qy	ACAGCTCTGCAGAACCTCATCTCTTCTTGGAGCAGCAGCTGGGTAGAGGTCTCTGGG	1962
1381	Db	ACAGCTCTGCAGAACCTCATCTCTTCTTGGAGCAGCAGCTGGGTAGAGGTCTCTGGG	1440
1963	Qy	TTCTCTGAGTCCAGCAGGAAGAAATATTTCTACAGATATTTCAAGATGAAAGCAAGCA	2022
1441	Db	TTCTCTGAGTCCAGCAGGAAGAAATATTTCTACAGATATTTCAAGATGAAAGCAAGCA	1500
2023	Qy	ATTAGAGCTTTTAGGTTGGTCAAAATCAACAAAGAGCTCTGGGCCCTGTCTTGTGCC	2082
1501	Db	ATTAGAGCTTTTAGGTTGGTCAAAATCAACAAAGAGCTCTGGGCCCTGTCTTGTGCC	1560
2083	Qy	TGGGTCTCTGGCTGGCTTGCACTTGGCTGTGACAGCAGATGAAAGCGGAAGAAAATC	2142
1561	Db	TGGGTCTCTGGCTGGCTTGCACTTGGCTGTGACAGCAGATGAAAGCGGAAGAAAATC	1620
2143	Qy	ACACTGACTTCCAAGACCAACCAACCTCTGTCTACATTAACCTTGCCAGGCTCTCCAA	2202
1621	Db	ACACTGACTTCCAAGACCAACCAACCTCTGTCTACATTAACCTTGCCAGGCTCTCCAA	1680
2203	Qy	GCTCAGCATTGGGACCCAGCTCAGACCTCTGCTCTGCTGGCTGTGAGGGCATCTGG	2262
1681	Db	GCTCAGCATTGGGACCCAGCTCAGACCTCTGCTCTGCTGGCTGTGAGGGCATCTGG	1740
2263	Qy	CAAAAAAGACCTTTTACGTCAGATGACCTCAGAAAGCATGGGTAGATGGGGCCATC	2322
1741	Db	CAAAAAAGACCTTTTACGTCAGATGACCTCAGAAAGCATGGGTAGATGGGGCCATC	1800
2323	Qy	ATCTCACTCTTGAAGATGGGTATTTCTCAAGAGCAACCCATCCTCTGAGCTACAGC	2382
1801	Db	ATCTCACTCTTGAAGATGGGTATTTCTCAAGAGCAACCCATCCTCTGAGCTACAGC	1860
2383	Qy	TTCACTTCACTCTTGAAGTCTTTGCAAGTCTTGGCAGCAATGCTCTATGCTTGGAGTGAG	2442
1861	Db	TTCACTTCACTCTTGAAGTCTTTGCAAGTCTTTGCAAGTCTTGGCAGCAATGCTCTATGCTTGGAGTGAG	1920
2443	Qy	AAGGGAGAGGTAAACAATTTGCAATTAAGATTTGGAAGAGCAAGCTAGAGCATAT	2502
1921	Db	AAGGGAGAGGTAAACAATTTGCAATTAAGATTTGGAAGAGCAAGCTAGAGCATAT	1980
2503	Qy	GGAATACATGGCTGTTGGGCACTCAACACAGCTTCTTATTTGGGCTGTTAAGTGAT	2562
1981	Db	GGAATACATGGCTGTTGGGCACTCAACACAGCTTCTTATTTGGGCTGTTAAGTGAT	2040
2563	Qy	GAGGGGAGAGAGATGGAGAACATCTTCACTGCGGCTGTCTCAGGGGAGAACCTG	2622
2041	Db	GAGGGGAGAGAGATGGAGAACATCTTCACTGCGGCTGTCTCAGGGGAGAACCTG	2100
2623	Qy	ATGAGTGGGTCCCGTCCCTGAGCTGTGCTGAGCCACACTCTCTGGAGTCCCTCCAC	2682
2101	Db	ATGAGTGGGTCCCGTCCCTGAGCTGTGCTGAGCCACACTCTCTGGAGTCCCTCCAC	2160
2683	Qy	TGCTTGTACGAGCTCGGAACAAAACGTTTCTCTGACAAAGTGATGGGCCAATTTCCGAAGA	2742
2161	Db	TGCTTGTACGAGCTCGGAACAAAACGTTTCTCTGACAAAGTGATGGGCCAATTTCCGAAGA	2222
2743	Qy	ATGGGCATGTGTAGAAAACAGACATGGAGCTCTTAGTGTGCACTTTTCTGCAATTAATTC	2802
2221	Db	ATGGGCATGTGTAGAAAACAGACATGGAGCTCTTAGTGTGCACTTTTCTGCAATTAATTC	2280
2803	Qy	AGCCGCCACGTGAAGAGCTTCACTGATTTGAGGGCAGGCGACAGATCAACATGGAGC	2862
2281	Db	AGCCGCCACGTGAAGAGCTTCACTGATTTGAGGGCAGGCGACAGATCAACATGGAGC	2340
2863	Qy	CCACCATGGTGTGCTTGTTCAGGTGGTCCCACTGATGCTTATTTGGCAGATTTCTC	2922
2341	Db	CCACCATGGTGTGCTTGTTCAGGTGGTCCCACTGATGCTTATTTGGCAGATTTCTC	2400
2923	Qy	TTCTCCGTCCTCAAGCTCACAGAAACCTGAAGAGCTGAGACCTAAAGTGAAACTCGCTG	2982
2401	Db	TTCTCCGTCCTCAAGCTCACAGAAACCTGAAGAGCTGAGACCTAAAGTGAAACTCGCTG	2460
2983	Qy	AGCCACTCTGCAGTGAAGAGCTTTTGAAGCCTTGAGACCGCCTCGCTGCTTCTGGAG	3042
2461	Db	AGCCACTCTGCAGTGAAGAGCTTTTGAAGCCTTGAGACCGCCTCGCTGCTTCTGGAG	2520
3043	Qy	ACCCTGGGTTGGCTGGCTCTCAGCTGAGGACTGCAAGGACCTTGGCTTGGG	3102
2521	Db	ACCCTGGGTTGGCTGGCTCTCAGCTGAGGACTGCAAGGACCTTGGCTTGGG	2580
3103	Qy	CTGAGAGCAACACAGACCTTGACCGAGCTGAGCTTCAATGTGCTCAGCGATGCT	3162
2581	Db	CTGAGAGCAACACAGACCTTGACCGAGCTGAGCTTCAATGTGCTCAGCGATGCT	2640
3163	Qy	GGAGCCAAACCTTTGCCAGAGCTGAGACAGCCGAGCTGCAAGCTACAGGACCTGCAG	3222
2641	Db	GGAGCCAAACCTTTGCCAGAGCTGAGACAGCCGAGCTGCAAGCTACAGGACCTGCAG	2700
3223	Qy	CTGCTCAGCTGTGGCTCTACGCTGTGACTGTGCTCAGGACCTGGGCTCTGTGCTTAGTGCC	3282
2701	Db	CTGCTCAGCTGTGGCTCTACGCTGTGACTGTGCTCAGGACCTGGGCTCTGTGCTTAGTGCC	2760
3283	Qy	AGCCCAAGCTTGAAGGAGCTAGACCTGCAGCAGAAACAACTGATGACGTTGGCGTGCGA	3342
2761	Db	AGCCCAAGCTTGAAGGAGCTAGACCTGCAGCAGAAACAACTGATGACGTTGGCGTGCGA	2820
3343	Qy	CTGCTCTGTGAGGGCTCAGGCACTCTGCTGCTGCAAACTCATACGCTGGGGCTGGACCAAG	3402
2821	Db	CTGCTCTGTGAGGGCTCAGGCACTCTGCTGCTGCAAACTCATACGCTGGGGCTGGACCAAG	2880
3403	Qy	ACAACCTGAGTGTAGATGAGAGGCAAGCTGAGGCTCTGAGCAGGAGAAACCTCAG	3462
2881	Db	ACAACCTGAGTGTAGATGAGAGGCAAGCTGAGGCTCTGAGCAGGAGAAACCTCAG	2940
3463	Qy	CTGCTCATCTTTCAGCAGACGAAACCAAGTGTGATGACCCCTACTCAGGGCTGGGATACG	3522
2941	Db	CTGCTCATCTTTCAGCAGACGAAACCAAGTGTGATGACCCCTACTCAGGGCTGGGATACG	3000
3523	Qy	GGAGAGTGTAGTAAATAGCACTCTCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGGCG	3582
3001	Db	GGAGAGTGTAGTAAATAGCACTCTCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGGCG	3060
3583	Qy	GCTTCCCATGTGTCTCAGGCTTAATCTCAAACTCTGGAAGTGTGAGTCTTCTCCCAAT	3642
3061	Db	GCTTCCCATGTGTGTCTCAGGCTTAATCTCAAACTCTCTGGAAGTGTGAGTCTTCTCCCAAT	3120
3643	Qy	GCTGAGATTTCAGAGGAAAGCTCCCGAGAGTGTAGTACCGGTGGAACTCTTGTGCTGCTC	3702
3121	Db	GCTGAGATTTCAGAGGAAAGCTCCCGAGAGTGTAGTACCGGTGGAACTCTTGTGCTGCTC	3180
3703	Qy	TCTCTGCTCTCAAGGGGACCTTSCATACGAAGCTTTTGGGGACTGACGATGACTTCTCG	3762
3181	Db	TCTCTGCTCTCAAGGGGACCTTSCATACGAAGCTTTTGGGGACTGACGATGACTTCTCG	3240
3763	Qy	GGCCCCACGGGCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTTGTACCGAGTT	3822
3241	Db	GGCCCCACGGGCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTTGTACCGAGTT	3300

Db 2821 |||||CTGCTCTGTGAGGGCTCAGGCATCTGCTCTGCAAACTCATACGCGCTGG-----2869
QY 3403 ACAACTCTGAGTGATGAGATGAGCGAGAACTGAGGGCCCTGGAGGAGAAACCTCAG 3462
Db 2870 -----2869
QY 3463 CTGCTCATCTTTCAGCAGACGGAAACCAAGTGTGATGACCCCTCTAGAGGGCCTGGATAGC 3522
Db 2870 -----GGAACCAAGTGTGATGACCCCTACTGAGGGCCTGGATAGC 2910
QY 3523 GGAGAGATGAGTAATAGCATCTCTCACTCAAGCGGCAGAGACTCGGATCAGAGGGCG 3582
Db 2911 GGAGAGATGAGTAATAGCATCTCTCACTCAAGCGGCAGAGACTCGGATCAGAGGGCG 2970
QY 3583 GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCTGGACGTGAGCAAGATCTTCCCAAT 3642
Db 2971 GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCTGGACGTGAGCAAGATCTTCCCAAT 3030
QY 3643 GCTGAGATTCAGAGAAAGCTCCCGAGAGGTAGTACCGGTGGAACCTCTGTGCGTGCCT 3702
Db 3031 GCTGAGATTCAGAGAAAGCTCCCGAGAGGTAGTACCGGTGGAACCTCTGTGCGTGCCT 3090
QY 3703 TCTCTGCTCTCAAGGGACCTGCAATACGAAGCCTTTGGGACTGACGATGACTTCTGG 3762
Db 3091 TCTCTGCTCTCAAGGGACCTGCAATACGAAGCCTTTGGGACTGACGATGACTTCTGG 3150
QY 3763 GGCCCAACGGGGCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTT 3822
Db 3151 GGCCCAACGGGGCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTT 3210
QY 3823 CACTTCCCTGTAGTGGCTCTACCGCTGGCCCAACACACGGGTCTCTGTCTTTGTGATGAGA 3882
Db 3211 CACTTCCCTGTAGTGGCTCTACCGCTGGCCCAACACACGGGTCTCTGTCTTTGTGATGAGA 3270
QY 3883 GAAGCGGTGACCGTTGAGATTGAATTTCTGTGTGGGAACAGTTCTGGGTGAGATCAAC 3942
Db 3271 GAAGCGGTGACCGTTGAGATTGAATTTCTGTGTGGGAACAGTTCTGGGTGAGATCAAC 3330
QY 3943 CACACACAGCTGGAATGTGGCAGGGCCTCTGCTGGACATCAAGGCTCAGCGCTGGAGCT 4002
Db 3331 CACACACAGCTGGAATGTGGCAGGGCCTCTGCTGGACATCAAGGCTCAGCGCTGGAGCT 3390
QY 4003 GTGGAGCTGTGCACCTCCCTCACTTTGTGGCTCTCCAAAGGGGGCCATGTGACACATCC 4062
Db 3391 GTGGAGCTGTGCACCTCCCTCACTTTGTGGCTCTCCAAAGGGGGCCATGTGACACATCC 3450
QY 4063 CTGTTCCAAATGGCCCACTTTAAAGAGGAGGGGATGCTCTGGAGAACCGCAGCGGTG 4122
Db 3451 CTGTTCCAAATGGCCCACTTTAAAGAGGAGGGGATGCTCTGGAGAACCGCAGCGGTG 3510
QY 4123 GAGCTGCATCAGATGTTCTGGAACCCAGCTTCTCCCTTTGGAGTCTCCTCTGAAA 4182
Db 3511 GAGCTGCATCAGATGTTCTGGAACCCAGCTTCTCCCTTTGGAGTCTCCTCTGAAA 3570
QY 4183 ATGATCCATAATGCTCGCTTCACTTCCCGTCACTCTGTGTGTGTGTTTACACACCGC 4242
Db 3571 ATGATCCATAATGCTCGCTTCACTTCCCGTCACTCTGTGTGTGTGTTTACACACCGC 3630
QY 4243 GTCATCTCTGAGGAAGTCACTTCCACTTACCTGATCCCAAGTGACTGCTCCATTGG 4302
Db 3631 GTCATCTCTGAGGAAGTCACTTCCACTTACCTGATCCCAAGTGACTGCTCCATTGG 3690
QY 4303 AAGGAATGGAGCTCTGCTATCGAAGCCTGGAGAAACAGAGTGTCTCGGAGTTCTAC 4362
Db 3691 AAGGAATGGAGCTCTGCTATCGAAGCCTGGAGAAACAGAGTGTCTCGGAGTTCTAC 3750
QY 4363 GTTGGCCACTTGGATTCAGGGATCAGCTGCAAGTGAAGACAAAGAAAGATGAGACTCTG 4422
Db 3751 GTTGGCCACTTGGATTCAGGGATCAGCTGCAAGTGAAGACAAAGAAAGATGAGACTCTG 3810
QY 4423 GTGTGGAGGGCCTTGTGTGAACACAGGAGATCTCATGCTGTCAACTACTCTGATCCCTCCA 4482
|||||

Db 3811 GTGTGGAGGGCTTGGTGAAACCAGGAGATCTATGCTGCTGCAACTACTCTGATCCCTCA 3870
QY 4483 GCCCGCATAGCGGTACCTTTCACCTCTGGATGCCCGCAGTTGCTGCTACACTTTGTGACCA 4542
Db 3871 GCCCGCATAGCGGTACCTTTCACCTCTGGATGCCCGCAGTTGCTGCTACACTTTGTGACCA 3930
QY 4543 TATCGAGAGCTGATGATAGCCGAGTGAATCGGTGGAGGTGCTTGGACAAACTGAT 4602
Db 3931 TATCGAGAGCTGATGATAGCCGAGTGAATCGGTGGAGGTGCTTGGACAAACTGAT 3990
QY 4603 GGACAGGTGCTGAGCCAGGACGATACGAGGGGTGCTGGCTGAGAACACAGAGGCCAGC 4662
Db 3991 GGACAGGTGCTGAGCCAGGACGATACGAGGGGTGCTGGCTGAGAACACAGAGGCCAGC 4050
QY 4663 CAGATGCGGAAAGCTGTTTTCAGCTTTCAGCCAGTCTCTGGACCGGAAAGTGCATAATGACTC 4722
Db 4051 CAGATGCGGAAAGCTGTTTTCAGCTTTCAGCCAGTCTCTGGACCGGAAAGTGCATAATGACTC 4110
QY 4723 TACCAAGCCTGAAGGAGACCCATCTCCTCATTTATGGAATCTCTGGGAGAGGGCAGC 4782
Db 4111 TACCAAGCCTGAAGGAGACCCATCTCCTCATTTATGGAATCTCTGGGAGAGGGCAGC 4170
QY 4783 AAAAAAGGACTCTCTGCCACTCAGCAGCTGA 4812
Db 4171 AAAAAAGGACTCTCTGCCACTCAGCAGCTGA 4200

RESULT 10
US-09-388-221-5
; Sequence 5, Application US/09388221A
; Publication NO. US20020192643A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. US20020192643A1el Card Proteins Involved in Cell Death Regula
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388,221A
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 4332
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(4332)
US-09-388-221-5

Query Match 72.7%; Score 3958; DB 10; Length 4332;
Best Local Similarity 95.0%; Pred. No. 0;
Matches 4200; Conservative 0; Mismatches 0; Indels 22; Gaps 2;

QY 523 ATGGCTGGCGGAGCCTGGGGCCGCTGGCTGTTTACTTGGAGTTCCTGAGAGGAGGAG 582
Db 1 ATGGCTGGCGGAGCCTGGGGCCGCTGGCTGTTTACTTGGAGTTCCTGAGAGGAGGAG 60
QY 583 CTGAGAGGAGTTCAGCTTCTGCTCCCAATAAGCGCATCCAGAGGCTCTTCCGGGTGAG 642
Db 61 CTGAGAGGAGTTCAGCTTCTGCTCCCAATAAGCGCATCCAGAGGCTCTTCCGGGTGAG 120
QY 643 ACACCCGCTCAGCCAGAGAAAGAGTGGCATGGAGGTGGCCTCGTACCTGGTGGCTCAG 702
Db 121 ACACCCGCTCAGCCAGAGAAAGAGTGGCATGGAGGTGGCCTCGTACCTGGTGGCTCAG 180
QY 703 TATGGGAGACAGCGGGCCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGGCTGAGG 762
Db 181 TATGGGAGACAGCGGGCCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGGCTGAGG 240
QY 763 TCACCTGTGGCCCAAGCCAGGAAAGGGCAGGCCACTCTCCCTCATTCCTCCCTACAGCCCA 822
Db 241 TCACCTGTGGCCCAAGCCAGGAAAGGGCAGGCCACTCTCCCTCATTCCTCCCTACAGCCCA 300
QY 823 AGTGAACCCCACTGGGGTCTCCAGGCCAACCCACCTCCACCCAGTGTCTAATGCTGCGCTGG 882

|||||
Db 301 AGTGAACCCCACTGGGGTCTCCAGCCAAAGCCACCTCCAGCGAGTCTAATGCCCTGG 360
QY 883 ATCCATGAATGCGCGGGGGTGCAACCGAGGCTCAGAGAAGGGTTTGTAGACAGCTG 942
Db 361 ATCCATGAATGCGCGGGGGTGCAACCGAGGCTCAGAGAAGGGTTTGTAGACAGCTG 420
QY 943 CTTGACACATCTGGACGCGCTGGAGAGAAATCTCTGCTCTCCTCTACCAAGCTCTT 1002
Db 421 CTTGACACATCTGGACGCGCTGGAGAGAAATCTCTGCTCTCCTCTACCAAGCTCTT 480
QY 1003 CCAAGCTCCCGACCATGAGTCTCCAAGCCAGGAGTCAACCAAGCCGCCCAATCCACA 1062
Db 481 CCAAGCTCCCGACCATGAGTCTCCAAGCCAGGAGTCAACCAAGCCGCCCAATCCACA 540
QY 1063 GAGTCTGGGAGCTGGGGATCCCACTCAGCCCGAGCTAGCACCCAGAGAGAGGAG 1122
Db 541 GAGTCTGGGAGCTGGGGATCCCACTCAGCCCGAGCTAGCACCCAGAGAGAGGAG 600
QY 1123 GCTCTGGGACCCCAATGGCTCTGGATGAAACGTACAGGAATTTACTACACAGAAATCAGA 1182
Db 601 GCTCTGGGACCCCAATGGCTCTGGATGAAACGTACAGGAATTTACTACACAGAAATCAGA 660
QY 1183 GAAAGAGAGAGAGAGAAATCAGAGAAAGCGAGCCGCCATGGGCGAGCGGTGGTAGGAACG 1242
Db 661 GAAAGAGAGAGAGAGAAATCAGAGAAAGCGAGCGCCGCCATGGGCGAGCGGTGGTAGGAACG 720
QY 1243 CCCCCAGCGGACACGAGCTACAGCCCGACACACCCATGGGAGCTTCTGTGAGA 1302
Db 721 CCCCCAGCGGACACGAGCTACAGCCCGACACACCCATGGGAGCTTCTGTGAGA 780
QY 1303 GAGAGCTCTGTTCCACATGGCCCTGGAAAAATGAGATTTTAAACAAAAATTCACACAG 1362
Db 781 GAGAGCTCTGTTCCACATGGCCCTGGAAAAATGAGATTTTAAACAAAAATTCACACAG 840
QY 1363 CTGCTACTTCTCAAAAGACTCACCCAGAGCCAGAGTCCCTGTCTCAAGAGAAGCTGG 1422
Db 841 CTGCTACTTCTCAAAAGACTCACCCAGAGCCAGAGTCCCTGTCTCAAGAGAAGCTGG 900
QY 1423 CTGATATTGTGGAGAGAAATCGAGACATTTAAATTGAGATCAGAGACTTATTGGCCCA 1482
Db 901 CTGATATTGTGGAGAGAAATCGAGACATTTAAATTGAGATCAGAGACTTATTGGCCCA 960
QY 1483 GGCCTGGATACCAAGAACTCGCATAGTCACTACGAGGGGCTCTCGAATTTGGGAG 1542
Db 961 GGCCTGGATACCAAGAACTCGCATAGTCACTACGAGGGGCTCTCGAATTTGGGAG 1020
QY 1543 TCAACACTGGCCAGGCTGAAGGAAGCTGGGGGAGAGCCAGCTGTATGGGACCGC 1602
Db 1021 TCAACACTGGCCAGGCTGAAGGAAGCTGGGGGAGAGCCAGCTGTATGGGACCGC 1080
QY 1603 TTCAGACTGTCTTACTTTCAGTCTCAGAGAGCTGGCCAGTCCAAAGTGTGTAGTCTC 1662
Db 1081 TTCAGACTGTCTTACTTTCAGTCTCAGAGAGCTGGCCAGTCCAAAGTGTGTAGTCTC 1140
QY 1663 GCTGAGCTCATCGGAAAGATGGGACAGCCACTCCGGCTCCCATAGACAGATCCTGTCT 1722
Db 1141 GCTGAGCTCATCGGAAAGATGGGACAGCCACTCCGGCTCCCATAGACAGATCCTGTCT 1200
QY 1723 AGGCCAGAGCGGTCTCTTCTCTCTCGATGGTGTAGATGAGCCAGATGGGTCTTGCA 1782
Db 1201 AGGCCAGAGCGGTCTCTTCTCTCTCGATGGTGTAGATGAGCCAGATGGGTCTTGCA 1260
QY 1783 GAGCCGAGTTCTGAGCTCTGTCTGCACTGGAGCCAGCCACAGCCGCGGATGCACTGTCTG 1842
Db 1261 GAGCCGAGTTCTGAGCTCTGTCTGCACTGGAGCCAGCCACAGCCGCGGATGCACTGTCTG 1320
QY 1843 GGCAGTTCTCGGGAAACTACTTCCGAGGCACTCTTCTGATCAGGCTCGGACC 1902
Db 1321 GGCAGTTCTCGGGAAACTACTTCCGAGGCACTCTTCTGATCAGGCTCGGACC 1380
QY 1903 ACAGTCTCGAAGAACTCAATTCCTTTGGAGCAGGCACTGTGGGTAGAGTCTCTGGGG 1962
|||||

Db 1381 ACAGCTCTGCAGAACTCAATTCCTTTTGGAGCAGGCACTGTGGGTAGAGTCTCTGGGG 1440
QY 1963 TTCTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTACAGATGAAAGGCAAGCA 2022
Db 1441 TTCTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTACAGATGAAAGGCAAGCA 1500
QY 2023 ATTAGAGCTTTAGGTTGGTCAAAATCAAAAGAGCTCTGGGCCCTGTCTTCTGTGCC 2082
Db 1501 ATTAGAGCTTTAGGTTGGTCAAAATCAAAAGAGCTCTGGGCCCTGTCTTCTGTGCC 1560
QY 2083 TGGGTGTCTGGCTGGCCCTGCACTTGCCTGATGACAGAGATGAAGCGGAAGAAAACTC 2142
Db 1561 TGGGTGTCTGGCTGGCCCTGCACTTGCCTGATGACAGAGATGAAGCGGAAGAAAACTC 1620
QY 2143 ACACGTGACTTCCAAGACCACCAACCCCTCTGTCTACATTTACCTTTGCCAGGCTCTCAA 2202
Db 1621 ACACGTGACTTCCAAGACCACCAACCCCTCTGTCTACATTTACCTTTGCCAGGCTCTCAA 1680
QY 2203 GCTCAGCCATTGGGACCCAGCTCAGAGACCTCTGCTCTCTGGCTGCTGAGGGCATCTGG 2262
Db 1681 GCTCAGCCATTGGGACCCAGCTCAGAGACCTCTGCTCTCTGGCTGCTGAGGGCATCTGG 1740
QY 2263 CAAAAAGAGACCTTTTTCAGTCCAGATGACCTCAGGAAGCATGGGTTAGATGGGGCCATC 2322
Db 1741 CAAAAAGAGACCTTTTTCAGTCCAGATGACCTCAGGAAGCATGGGTTAGATGGGGCCATC 1800
QY 2323 ATCTCCACTTCTTGAAGTGGGTATTTCTCAAGAGCACCCATCCCTCTGAGCTACAGC 2382
Db 1801 ATCTCCACTTCTTGAAGTGGGTATTTCTCAAGAGCACCCATCCCTCTGAGCTACAGC 1860
QY 2383 TTCAATTCACCTCTGTTTCCAAAGAGTCTTTTTCAGCAATGTCTCTATGTCTTGAGGATGAG 2442
Db 1861 TTCAATTCACCTCTGTTTCCAAAGAGTCTTTTTCAGCAATGTCTCTATGTCTTGAGGATGAG 1920
QY 2443 AAGGGAGAGGTAAACATTTCTAATTCATATGATTTGGAAGACGCTGAAGACATAT 2502
Db 1921 AAGGGAGAGGTAAACATTTCTAATTCATATGATTTGGAAGACGCTGAAGACATAT 1980
QY 2503 GGAATACATGGCTGTTTGGGGCATCAACCAACGTTTCTTATTTGGGCTGTGTAAGTGTAT 2562
Db 1981 GGAATACATGGCTGTTTGGGGCATCAACCAACGTTTCTTATTTGGGCTGTGTAAGTGTAT 2040
QY 2563 GAGGGGAGAGAGATGGAGAAACATCTTTCACTGCGGCTGTCTCAGGGAGGAAACCTG 2622
Db 2041 GAGGGGAGAGAGATGGAGAAACATCTTTCACTGCGGCTGTCTCAGGGAGGAAACCTG 2100
QY 2623 ATGCAGTGGTCCCGTCCCTGCGAGCTGTGTGTGACGACACTCTCTGGAGTCCCTCCAC 2682
Db 2101 ATGCAGTGGTCCCGTCCCTGCGAGCTGTGTGTGACGACACTCTCTGGAGTCCCTCCAC 2160
QY 2683 TGCTTGTACGAGACTCGGAACAAACGTTTCTGTGACAAAGTGAAGTGGCCCATTTGGAAGAA 2742
Db 2161 TGCTTGTACGAGACTCGGAACAAACGTTTCTGTGACAAAGTGAAGTGGCCCATTTGGAAGAA 2220
QY 2743 ATGGSCATGTGTATAGAAACAGACATGAGAGCTCTTAGTGTGCACTTTCTGCAATTAATTC 2802
Db 2221 ATGGSCATGTGTATAGAAACAGACATGAGAGCTCTTAGTGTGCACTTTCTGCAATTAATTC 2280
QY 2803 AGCCGCCACGTAAGAGCTTTCAGCTGATTTGAGGSCAGGACGACAGATCAACATGGAGC 2862
Db 2281 AGCCGCCACGTAAGAGCTTTCAGCTGATTTGAGGSCAGGACGACAGATCAACATGGAGC 2340
QY 2863 CCCACATGGTAGTCTCTTTCAGGTGGGTCCAGTCAAGATGCCATTTGSCAGATTTCTC 2922
Db 2341 CCCACATGGTAGTCTCTTTCAGGTGGGTCCAGTCAAGATGCCATTTGSCAGATTTCTC 2400
QY 2923 TTCTCCGTCTCAAGGTCAACAGAAAACCTGAAGAGCTGGACCTAAGTGGAAAACCTCGCTG 2982
Db 2401 TTCTCCGTCTCAAGGTCAACAGAAAACCTGAAGAGCTGGACCTAAGTGGAAAACCTCGCTG 2460
QY 2983 AGCCACTCTGAGTGAAGAGTCTTTGTAAAGCCCTTGAGACGCGCTCTCGTCTCTGGAG 3042
Db 2461 AGCCACTCTGAGTGAAGAGTCTTTGTAAAGCCCTTGAGACGCGCTCTCGTCTCTGGAG 2520
|||||

QY 3043 ACCCTCGGTTGGCTGGCTGTGGCTCAAGCTGAGGACTGCAAGGACCTTGCCTTTGGG 3102
 Db 2521 ACCCTCGGTTGGCTGGCTGTGGCTCAAGCTGAGGACTGCAAGGACCTTGCCTTTGGG 2580
 QY 3103 CTGAGAGCCAAACAGACCTTGCACCGAGCTGGACCTGAGCTTCAATGTGCTCAGGATGCT 3162
 Db 2581 CTGAGAGCCAAACAGACCTTGCACCGAGCTGGACCTGAGCTTCAATGTGCTCAGGATGCT 2640
 QY 3163 GGAGCCAAACACCTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGGAGCTGACG 3222
 Db 2641 GGAGCCAAACACCTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGGAGCTGACG 2700
 QY 3223 CTGGTCAGTGTGGCTCAAGCTGTGACTGTGTCAGGACCTGGCTCTGTGCTTGTAGTGC 3282
 Db 2701 CTGGTCAGTGTGGCTCAAGCTGTGACTGTGTCAGGACCTGGCTCTGTGCTTGTAGTGC 2760
 QY 3283 AGCCCCAGCTGAGAGCTAGACCTGCGACGAGAA CAACCTGATGACGCTTGGCGTGGCA 3342
 Db 2761 AGCCCCAGCTGAGAGCTAGACCTGCGACGAGAA CAACCTGATGACGCTTGGCGTGGCA 2820
 QY 3343 CTGCTCTGTGAGGGGCTCAGGATCTGCTGCAAACTCATACGCTGGGGCTGGACACG 3402
 Db 2821 CTGCTCTGTGAGGGGCTCAGGATCTGCTGCAAACTCATACGCTGG----- 2869
 QY 3403 ACAACTCTGAGTGATGAGTAGGCGAGGAA CTGAGGGCCCTGGAGCAGGAGAACTCAG 3462
 Db 2870 ----- 2869
 QY 3463 CTGCTCATCTTCAGCAGACGGAACCAAGTGATGATGACCCCTACTGAGGGCCTGGATAG 3522
 Db 2870 -----GGAAACCAAGTGATGATGACCCCTACTGAGGGCCTGGATAG 2910
 QY 3523 GGAGAGATGAGTAATAGCATCTCTCACTCAAGCGCAGAGACTCGGATCAGAGAGGGG 3582
 Db 2911 GGAGAGATGAGTAATAGCATCTCTCACTCAAGCGCAGAGACTCGGATCAGAGAGGGG 2970
 QY 3583 GCTTCCCATGTTGCTCAGGCTAACTCTCAAACTCTGAGAGCTGAGCAAGATCTTCCCAATT 3642
 Db 2971 GCTTCCCATGTTGCTCAGGCTAACTCTCAAACTCTGAGAGCTGAGCAAGATCTTCCCAATT 3030
 QY 3643 GCTGAGATGAGAGAGAAAGCTCCCAAGAGTAGTACCGGTGAACTCTTGTGCGTGCCT 3702
 Db 3031 GCTGAGATGAGAGAGAAAGCTCCCAAGAGTAGTACCGGTGAACTCTTGTGCGTGCCT 3090
 QY 3703 TCTCTGCTCTCAAGGGGACCTGATACGAGGCTTTGGGACTCAGCATGACTTCTCG 3762
 Db 3091 TCTCTGCTCTCAAGGGGACCTGATACGAGGCTTTGGGACTCAGCATGACTTCTCG 3150
 QY 3763 GGCCCCCAGCGGCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTT 3822
 Db 3151 GGCCCCCAGCGGCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTT 3210
 QY 3823 CACTTCCCTGTAGCTGGCTCTACCGGTGCCCCAA CACGGGTCTCTGCTTTGTGATGAGA 3882
 Db 3211 CACTTCCCTGTAGCTGGCTCTACCGGTGCCCCAA CACGGGTCTCTGCTTTGTGATGAGA 3270
 QY 3883 GAGCGGTGACCGCTGAGATTGAAATTCTGTGTGTGGGACAGTTCTGGTGGATGATCAAC 3942
 Db 3271 GAGCGGTGACCGCTGAGATTGAAATTCTGTGTGTGGGACAGTTCTGGTGGATGATCAAC 3330
 QY 3943 CCACAGCAGCTGGATGAGTGGCAGGGCCCTCTGCTGGACATCAAGGCTCAGCGCTGAGCT 4002
 Db 3331 CCACAGCAGCTGGATGAGTGGCAGGGCCCTCTGCTGGACATCAAGGCTCAGCGCTGAGCT 3390
 QY 4003 GTGGAAAGTGTGACCTCCCTCACTTTGTGGCTCTCCAGGGGGCCCATGTGACACATCC 4062
 Db 3391 GTGGAAAGTGTGACCTCCCTCACTTTGTGGCTCTCCAGGGGGCCCATGTGACACATCC 3450
 QY 4063 CTGTTCCAAATGSCCCCACTTTAAAGAGAGGGGATGCTCTGTGAGAGCCAGCGGGTG 4122
 Db 3451 CTGTTCCAAATGSCCCCACTTTAAAGAGAGGGGATGCTCTGTGAGAGCCAGCGGGTG 3510

QY 4123 GAGCTGCATCATATAGTTCTTGGAACCCCGAGCTTCTCCCTTGGGAGTCTCTCTGAAA 4182
 Db 3511 GAGCTGCATCATATAGTTCTTGGAACCCCGAGCTTCTCCCTTGGGAGTCTCTCTGAAA 3570
 QY 4183 ATGATCCCAATAAGCCCTGCTTCAATCCCGTCACTCTGTGGTGTGTTTACCAACCG 4242
 Db 3571 ATGATCCCAATAAGCCCTGCTTCAATCCCGTCACTCTGTGGTGTGTTTACCAACCG 3630
 QY 4243 GTCCATCTCTGAGGAAGTCACTTCCACCTCTACCTGATCCCAAGTGTCTTCAATCGG 4302
 Db 3631 GTCCATCTCTGAGGAAGTCACTTCCACCTCTACCTGATCCCAAGTGTCTTCAATCGG 3690
 QY 4303 ----- 4302
 Db 3691 AAGGCCATAGATGATCTAGAAATGAAATTCCAGTTTGTGCGAATCCA CAAGCCACCCCG 3750
 QY 4303 ----- 4302
 Db 3751 CTGACCCCACTTTATATAGGCTGTCTTACACTGTGTCTGGGTCTGGTTCAGGGATGCTG 3810
 QY 4303 -----AAGGAACTGGAGCTCTGCTATCGAAGCCCTGAGAAAGACCAAGCTGTTTC 4350
 Db 3811 GAAATACTCCCAAGGAACTGGAGCTCTGCTATCGAAGCCCTGAGAAAGACCAAGCTGTTTC 3870
 QY 4351 TCGGAGTTCTAGCTTGGCCACTTGGGATCAGGGATCAGGCTGCAAGTGAAAGACCAAGAA 4410
 Db 3871 TCGGAGTTCTAGCTTGGCCACTTGGGATCAGGGATCAGGCTGCAAGTGAAAGACCAAGAA 3930
 QY 4411 GATGAGACTCTGTGTGGGAGGCTTGGTGAACAGGAGATCTCATGCTGCAACTACT 4470
 Db 3931 GATGAGACTCTGTGTGGGAGGCTTGGTGAACAGGAGATCTCATGCTGCAACTACT 3990
 QY 4471 CTGATCCCTCCAGCCCGCATAGCCGTACCTTCACTCTGATGATGCCCGCAGTTGCTGCAC 4530
 Db 3991 CTGATCCCTCCAGCCCGCATAGCCGTACCTTCACTCTGATGATGCCCGCAGTTGCTGCAC 4050
 QY 4531 TTTGTGGACCACTGAGAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 4590
 Db 4051 TTTGTGGACCACTGAGAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 4110
 QY 4591 GACAACTGCATGAGCAGGCTGCTGAGCCAGGAGCAGTACGAGAGGCTGCTGCTGAGAAC 4650
 Db 4111 GACAACTGCATGAGCAGGCTGCTGAGCCAGGAGCAGTACGAGAGGCTGCTGCTGAGAAC 4170
 QY 4651 ACAGGCCCGCAGCAGATGCGGAACTGTTTCACTTGAAGCAGCTCTGAGCCGGAAGTGC 4710
 Db 4171 ACAGGCCCGCAGCAGATGCGGAACTGTTTCACTTGAAGCAGCTCTGAGCCGGAAGTGC 4230
 QY 4711 AAAGATGAGCTCTCAAGCCCTGAAAGGAGACCCATCTCCTCATTTATGGAATCTTGG 4770
 Db 4231 AAAGATGAGCTCTCAAGCCCTGAAAGGAGACCCATCTCCTCATTTATGGAATCTTGG 4290
 QY 4771 GAGAGGGCAGCAAAAGGGACTCTCTGCCACTCAGCAGCTGA 4812
 Db 4291 GAGAGGGCAGCAAAAGGGACTCTCTGCCACTCAGCAGCTGA 4332

RESULT 11

US-09-388-221-9
 ; Sequence 9, Application US/09388221A
 ; Publication NO. US20020192643A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, John C.
 ; TITLE OF INVENTION: No. US20020192643A1el Card Proteins Involved in Cell Death Regul
 ; FILE REFERENCE: P-LJ 3650
 ; CURRENT APPLICATION NUMBER: US/09/388,221A
 ; CURRENT FILING DATE: 1999-09-01
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 9
 ; LENGTH: 4556
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence


```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) (4362)
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Construct
; US-09-388-221-9

Query Match      62.5%; Score 3400.4; DB 10; Length 4556;
Best Local Similarity 93.7%; Pred. No. 0;
Matches 3555; Conservative 0; Mismatches 236; Indels 3; Gaps 1;

QY 523 ATGGCTGGGGAGCCCTGGGGCCGCTGTTACTTGGAGTTCTCGAAGAGGAGGAG 582
DB 1 ATGGCTGGGGAGCCCTGGGGCCGCTGTTACTTGGAGTTCTCGAAGAGGAGGAG 60
QY 583 CTGAAGAGTTCCAGCTTCTGTCGCCAATAAAGCGCACTCCAGAGCTCTTCGGGTGAG 642
DB 61 CTGAAGAGTTCCAGCTTCTGTCGCCAATAAAGCGCACTCCAGAGCTCTTCGGGTGAG 120
QY 643 ACACCCGCTCAGCCAGAGAGAGCTGCGATGGAGGTGGCTCGTACCTGGTGGCTCAG 702
DB 121 ACACCCGCTCAGCCAGAGAGAGCTGCGATGGAGGTGGCTCGTACCTGGTGGCTCAG 180
QY 703 TATGGGGAGCAGCGGGCCCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGGCTGAGG 762
DB 181 TATGGGGAGCAGCGGGCCCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGGCTGAGG 240
QY 763 TCACTGTGCGCCCAAGCCAGGAAGGGGAGGCCACTCTCCCTCATTCGCCCTACAGCCCA 822
DB 241 TCACTGTGCGCCCAAGCCAGGAAGGGGAGGCCACTCTCCCTCATTCGCCCTACAGCCCA 300
QY 823 AGTGAACCCCACTGGGGTCTCCAGCCCAACCCACCTCCACCGCAGTGTAAATGCCCTGG 882
DB 301 AGTGAACCCCACTGGGGTCTCCAGCCCAACCCACCTCCACCGCAGTGTAAATGCCCTGG 360
QY 883 ATCCATGAATTTGCCGGGGGGTGACCCAGGGCTCAGAGAGAAGGTTTGGACAGAGCTG 942
DB 361 ATCCATGAATTTGCCGGGGGGTGACCCAGGGCTCAGAGAGAAGGTTTGGACAGAGCTG 420
QY 943 CTGGAACAATCTGGAACCGCGCTGGAGAGAAATCTTGCCCTCACTCTCTACCAAGCTCTT 1002
DB 421 CTGGAACAATCTGGAACCGCGCTGGAGAGAAATCTTGCCCTCACTCTCTACCAAGCTCTT 480
QY 1003 CCAAGCTCCCCAGACCATGATCTCAAGCCAGGAGTCAACCAACCCGCCCATCCACA.1062
DB 481 CCAAGCTCCCCAGACCATGATCTCTCAAGCCAGGAGTCAACCAACCCGCCCATCCACA 540
QY 1063 GCAGTGCTGGGAGCTGGGGATCCCCACCTCAGCCCCAGCTAGCACCCAGAGAGCAGGAG 1122
DB 541 GCAGTGCTGGGAGCTGGGGATCCCCACCTCAGCCCCAGCTAGCACCCAGAGAGCAGGAG 600
QY 1123 GCTCTGGGACCAATGGGCTCTGGATGAAAGTCAAGAAATTTACTTACAGAAATCAGA 1182
DB 601 GCTCTGGGACCAATGGGCTCTGGATGAAAGTCAAGAAATTTACTTACAGAAATCAGA 660
QY 1183 GAAAGAGAGAGAGAGAAATCAGAGAAAGGAGGCCGCCATGGGCAGCGGTGGTAGAACG 1242
DB 661 GAAAGAGAGAGAGAGAAATCAGAGAAAGGAGGCCGCCATGGGCAGCGGTGGTAGAACG 720
QY 1243 CCCCCACAGCGCACACAGCTTACAGCCCCACCAACCCATGGGAGGCTTCTGTGAGA 1302
DB 721 CCCCCACAGCGCACACAGCTTACAGCCCCACCAACCCATGGGAGGCTTCTGTGAGA 780
QY 1303 GAGAGCTCTGTTCACATATGGGCCCTGGAAAAATGAGGATTTTAAACCAAAATTCACACAG 1362
DB 781 GAGAGCTCTGTTCACATATGGGCCCTGGAAAAATGAGGATTTTAAACCAAAATTCACACAG 840
QY 1363 CTGCTACTTCTACAAAGACTTACCCAGAGCCAGAGATCCCTGGTCAAGAGAGCTGG 1422
DB 841 CTGCTACTTCTACAAAGACTTACCCAGAGCCAGAGATCCCTGGTCAAGAGAGCTGG 900
QY 1423 CTTGATTTATGTGGAGAGAAATCAGAGACATTTAAATTTGAGATCAGAGACTTTATTTGGCCCA 1482
DB 1482 CTTGATTTATGTGGAGAGAAATCAGAGACATTTAAATTTGAGATCAGAGACTTTATTTGGCCCA 960
QY 1483 GGCCTGGATACCCAGAAACCTCGCATAGTCACTACTGCAAGGGGCTGTGGAAATTTGGGAG 1542
DB 961 GGCCTGGATACCCAGAAACCTCGCATAGTCACTACTGCAAGGGGCTGTGGAAATTTGGGAG 1020
QY 1543 TCAACACACTGGCCAGGCTGAAGGAAGCCTGGGGGAGAGGCCAGCTGTATGGGACCGC 1602
DB 1021 TCAACACACTGGCCAGGCTGAAGGAAGCCTGGGGGAGAGGCCAGCTGTATGGGACCGC 1080
QY 1603 TTCCAGCATGTCTTCTACTTCCAGCTGCAAGAGCTGGGCCAGTCCAAAGTGTGTAGTCTC 1662
DB 1081 TTCCAGCATGTCTTCTACTTCCAGCTGCAAGAGCTGGGCCAGTCCAAAGTGTGTAGTCTC 1140
QY 1663 GCTGAGCTCATCGGAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCCTGTCT 1722
DB 1141 GCTGAGCTCATCGGAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCCTGTCT 1200
QY 1723 AGGCCAGAGCGGCTCTCTTTCATCTCGATGGTGTAGATGAGCCAGGATGGGTCTTGCGAG 1782
DB 1201 AGGCCAGAGCGGCTCTCTTTCATCTCGATGGTGTAGATGAGCCAGGATGGGTCTTGCGAG 1260
QY 1783 GAGCCGAGTTCTGAGCTCTGTCTGCACTGGAGCCAGCCAGCCGCGGATGCACTGTCTG 1842
DB 1261 GAGCCGAGTTCTGAGCTCTGTCTGCACTGGAGCCAGCCAGCCGCGGATGCACTGTCTG 1320
QY 1843 GGCAGTTTGTGGGAAACTATACTTCCCGAGGATCCTTCCCTGATCAGGCTCGGACC 1902
DB 1321 GGCAGTTTGTGGGAAACTATACTTCCCGAGGATCCTTCCCTGATCAGGCTCGGACC 1380
QY 1903 ACAGCTCTGCGAAGCCTCAATCTCTTCTTGGAGCAGGACAGTCTGGGTAGAGTCTCTGGG 1962
DB 1381 ACAGCTCTGCGAAGCCTCAATCTCTTCTTGGAGCAGGACAGTCTGGGTAGAGTCTCTGGG 1440
QY 1963 TTCTCTGAGTCAGCAGGAAGAAATTTCTACAGATATTTCAAGATGAAAGGCAAGCA 2022
DB 1441 TTCTCTGAGTCAGCAGGAAGAAATTTCTACAGATATTTCAAGATGAAAGGCAAGCA 1500
QY 2023 ATTAGAGCTTTAGGTTGGTCAAAATCAAAAGAGCTCTGGGCCCTGTCTCTTGCC 2082
DB 1501 ATTAGAGCTTTAGGTTGGTCAAAATCAAAAGAGCTCTGGGCCCTGTCTCTTGCC 1560
QY 2083 TGGGTGTCTGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2142
DB 1561 TGGGTGTCTGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
QY 2143 ACACCTGACTTCAAGACCAACCAACCAACCCCTCTGTCTACATTTACCTTGTCCAGGCTCTCCAA 2202
DB 1621 ACACCTGACTTCAAGACCAACCAACCAACCCCTCTGTCTACATTTACCTTGTCCAGGCTCTCCAA 1680
QY 2203 GCTCAGCATTTGGGACCCAGCTCAGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 2262
DB 1681 GCTCAGCATTTGGGACCCAGCTCAGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 1740
QY 2263 CAAAAAAGAGCCCTTTTCAGTCCAGATGACCTCAGAAAGCATGGGTAGATGGGGCCATC 2322
DB 1741 CAAAAAAGAGCCCTTTTCAGTCCAGATGACCTCAGAAAGCATGGGTAGATGGGGCCATC 1800
QY 2323 ATCTCCACCTTTTGAAGATGGGTATTTCTTCAAGAGCAACCCATCCCTCTGAGCTACAGC 2382
DB 1801 ATCTCCACCTTTTGAAGATGGGTATTTCTTCAAGAGCAACCCATCCCTCTGAGCTACAGC 1860
QY 2383 TTCAATTCACCTCTGTTTCCAGAGTTCTTTTGAGCAATGTCTCTATGCTCTTGAGAGTATG 2442
DB 1861 TTCAATTCACCTCTGTTTCCAGAGTTCTTTTGAGCAATGTCTCTATGCTCTTGAGAGTATG 1920
QY 2443 AAGGGGAGAGGTAAACATTTCTAAATTTGCAATCATAGATTTGGAAAGACGCTAGAGCATAT 2502
DB 1921 AAGGGGAGAGGTAAACATTTCTAAATTTGCAATCATAGATTTGGAAAGACGCTAGAGCATAT 1980
QY 2503 GGAATATATGCGCTGTTTGGGGCATCAACCAACCGTTTCTTATTTGGGCTGTGTAAGTAT 2562
DB 1981 GGAATATATGCGCTGTTTGGGGCATCAACCAACCGTTTCTTATTTGGGCTGTGTAAGTAT 2040
```


Best Local Similarity 91.3%; Pred. No. 0; Matches 3465; Conservative 0; Mismatches 236; Indels 93; Gaps 2;

QY	523	ATGGCTGGCGGAGCTTGGGCGCGCTGGCTGTACTTGGAGTTCTGAAGAGGAG	582
Db	1	ATGGCTGGCGGAGCTTGGGCGCGCTGGCTGTACTTGGAGTTCTGAAGAGGAGGAG	60
QY	583	CTGAAGGAGTTCCAGCTTCTGCTCGCCAAATAAGCGCACTCCAGGAGCTTTCGGGTGAG	642
Db	61	CTGNAGAGTTCCAGCTTCTGCTCGCCAAATAAGCGCACTCCAGGAGCTTTCGGGTGAG	120
QY	643	ACACCCGCTCAGCCAGAGAACGAGTGGATGGAGGTGGCTCGTACCTGFGGCTCAG	702
Db	121	ACACCCGCTCAGCCAGAGAACGAGTGGATGGAGGTGGCTCGTACCTGFGGCTCAG	180
QY	703	TATGGGAGGAGCGGCGCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGCTGAGG	762
Db	181	TATGGGAGGAGCGGCGCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGCTGAGG	240
QY	763	TCACTGTGGCCCAAGCCAGGAAAGGCGAGGCACTCTCCCTCATCTCCCTACAGCCCA	822
Db	241	TCACTGTGGCCCAAGCCAGGAAAGGCGAGGCACTCTCCCTCATCTCCCTACAGCCCA	300
QY	823	AGTGAACCCCACTGGGGTCTCCAGCCAAACCACTCCACCGCAGTCTAATGCCCTGG	882
Db	301	AGTGAACCCCACTGGGGTCTCCAGCCAAACCACTCCACCGCAGTCTAATGCCCTGG	360
QY	883	ATCCATGAATTCGGCGGGGGTGACCCAGGGCTCAGAGAGAGGGTTTGGAGACAGTG	942
Db	361	ATCCATGAATTCGGCGGGGGTGACCCAGGGCTCAGAGAGAGGGTTTGGAGACAGTG	420
QY	943	CTTGACACATCTGGAGCGCGCTGGAGAGAAATCTCTGCTCTACCTCTTACCAAGCTCTT	1002
Db	421	CTTGACACATCTGGAGCGCGCTGGAGAGAAATCTCTGCTCTACCTCTTACCAAGCTCTT	480
QY	1003	CCAAGCTCCCCAGACCATGAGTCTCAAGCAGGAGTCAACCAAGCCGCCCAATCCACA	1062
Db	481	CCAAGCTCCCCAGACCATGAGTCTCAAGCAGGAGTCAACCAAGCCGCCCAATCCACA	540
QY	1063	GCAGTCTGGGAGTGGGATCCCACTCAGCCAGCGCTAGCACCCAGAGAGGAG	1122
Db	541	GCAGTCTGGGAGTGGGATCCCACTCAGCCAGCGCTAGCACCCAGAGAGGAG	600
QY	1123	GCTCTGGGACCAATGGCTCTGAGTGAACGCTCAGGAAATTTACTACAGAAATCAGA	1182
Db	601	GCTCTGGGACCAATGGCTCTGAGTGAACGCTCAGGAAATTTACTACAGAAATCAGA	560
QY	1183	GAAAGAGAGAGAGAAATCAGAGAAAGCGAGGCGCCCATGGGCGCGGTGGTAGGAA	1242
Db	661	GAAAGAGAGAGAGAAATCAGAGAAAGCGAGGCGCCCATGGGCGCGGTGGTAGGAA	720
QY	1243	CCCCACAGCGGACACAGAGCTACAGCCCGACCAACCACTGGGAGCTTCTGTGAGA	1302
Db	721	CCCCACAGCGGACACAGAGCTACAGCCCGACCAACCACTGGGAGCTTCTGTGAGA	780
QY	1303	GAGAGCTCTGTTCACATGGCCCTGGAAAAATGAGGATTTAAACAAAAATTCACACAG	1362
Db	781	GAGAGCTCTGTTCACATGGCCCTGGAAAAATGAGGATTTAAACAAAAATTCACACAG	840
QY	1363	CTGCTACTTCTTCAAAAGACCTACCCCGAGAGCGCAAGATCCCTGGTCAAGAGAGCTGG	1422
Db	841	CTGCTACTTCTTCAAAAGACCTACCCCGAGAGCGCAAGATCCCTGGTCAAGAGAGCTGG	900
QY	1423	CCTGATATGTGGAGAGAGATCAGAGACATTTAATTTAGATCAGAGATTTATGGCCCA	1482
Db	901	CCTGATATGTGGAGAGAGATCAGAGACATTTAATTTAGATCAGAGATTTATGGCCCA	960
QY	1483	GGCTCGATACCAAGAACTCGCATAGTATCATCTCAGGGGCTCTCTGAAATTTGGGAG	1542
Db	961	GGCTCGATACCAAGAACTCGCATAGTATCATCTCAGGGGCTCTCTGAAATTTGGGAG	1020
QY	1543	TCAACACTGGCAGGAGTGAAGAGCTTGGGGAGAGGCGCAGCTGTATGGGAGCCGC	1602
Db			

Db	1021	TCAACACTGGCCAGGCTGAAGGAAGCTGGGGAGAGGCCAGCTGTATGGGAGCCGC	1080
QY	1603	TTCCAGCAGATGCTTCTACTTTCAGCTGCAGAGAGCTGGGCCAGTCCAAAGTGGTGAAGTCTC	1662
Db	1081	TTCCAGCAGATGCTTCTACTTTCAGCTGCAGAGAGCTGGGCCAGTCCAAAGTGGTGAAGTCTC	1140
QY	1663	GCTGAGCTCATCGGAAAAAGATGGGACAGCACTCCGGCTCCCATTTAGACAGATCTGTCT	1722
Db	1141	GCTGAGCTCATCGGAAAAAGATGGGACAGCACTCCGGCTCCCATTTAGACAGATCTGTCT	1200
QY	1723	AGGCCAGAGCGGCTCTCTTCTCATCTCGATGGTGTAGATGAGCCAGGATGGGCTCTGAG	1782
Db	1201	AGGCCAGAGCGGCTCTCTTCTCATCTCGATGGTGTAGATGAGCCAGGATGGGCTCTGAG	1260
QY	1783	GAGCCGAGTTCTGAGCTCTGTCTGCACTGGAGCGAGCCACAGCCGGCGGATGCATGCTG	1842
Db	1261	GAGCCGAGTTCTGAGCTCTGTCTGCACTGGAGCGAGCCACAGCCGGCGGATGCATGCTG	1320
QY	1843	GGCAGTTTGTGGGAAAAATATACTTCCGAGGCGATCTTCTCTGATCACGGCTCGGACC	1902
Db	1321	GGCAGTTTGTGGGAAAAATATACTTCCGAGGCGATCTTCTCTGATCACGGCTCGGACC	1380
QY	1903	ACAGCTCTGCAGAACCTCATTTCTTTTGGAGCAGGCGAGCTTGGGTAGAGTCTCTGGG	1962
Db	1381	ACAGCTCTGCAGAACCTCATTTCTTTTGGAGCAGGCGAGCTTGGGTAGAGTCTCTGGG	1440
QY	1963	TTCTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTTACAGATATTTTACAGAT	2022
Db	1441	TTCTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTTACAGATATTTTACAGAT	1500
QY	2023	ATTAGAGCTTTTGTGGTCAAAATCAAAAGAGCTCTGGGCCCTGTGTCTTGTGCC	2082
Db	1501	ATTAGAGCTTTTGTGGTCAAAATCAAAAGAGCTCTGGGCCCTGTGTCTTGTGCC	1560
QY	2083	TGGGTGTCTGGCTGGCTGCACTTGTCTGATGAGCAGATGAAGCGGAGAAAGAACTC	2142
Db	1561	TGGGTGTCTGGCTGGCTGCACTTGTCTGATGAGCAGATGAAGCGGAGAAAGAACTC	1620
QY	2143	ACACTGACTTCCAGACCAACCAACCTCTGTCTACATTTACCTTGGCCAGGCTCTCAA	2202
Db	1621	ACACTGACTTCCAGACCAACCAACCTCTGTCTACATTTACCTTGGCCAGGCTCTCAA	1680
QY	2203	GCTCAGCATTGGGACCCAGCTCAGAGACCTCTGCTCTCTGGTGTCTGAGGGCATCTGG	2262
Db	1681	GCTCAGCATTGGGACCCAGCTCAGAGACCTCTGCTCTCTGGTGTCTGAGGGCATCTGG	1740
QY	2263	CAAAAAAGACCCCTTTTTCAGTCCAGATGACCTCAGGAAGCATGGGTTAGATGGGGCCATC	2322
Db	1741	CAAAAAAGACCCCTTTTTCAGTCCAGATGACCTCAGGAAGCATGGGTTAGATGGGGCCATC	1800
QY	2323	ATCTCCACCTTCTTGAAGATGGGTATTTCTTCAAGAGCACCCATCCCTCTGAGCTACAGC	2382
Db	1801	ATCTCCACCTTCTTGAAGATGGGTATTTCTTCAAGAGCACCCATCCCTCTGAGCTACAGC	1860
QY	2383	TTCAATTCACCTCTGTTTCCAAAGAGTTCTTTGAGCAATGTCTCTATGTCTTGGAGATGAG	2442
Db	1861	TTCAATTCACCTCTGTTTCCAAAGAGTTCTTTGAGCAATGTCTCTATGTCTTGGAGATGAG	1920
QY	2443	AAGGGAGAGGTAAACATTTCTAATTTGATATAGATTTGGAAAAAGACGCTAGAGCATAT	2502
Db	1921	AAGGGAGAGGTAAACATTTCTAATTTGATATAGATTTGGAAAAAGACGCTAGAGCATAT	1980
QY	2503	GGAAATACATGGCTCTTGGGGCATCAACCAAGCTTTCCTATTTGGGCTGTGTTAGTAT	2562
Db	1981	GGAAATACATGGCTCTTGGGGCATCAACCAAGCTTTCCTATTTGGGCTGTGTTAGTAT	2040
QY	2563	GAGGGGAGAGAGATGGAGAAACATTTTCACTCCGGCTGTCTCAGGGGAGGAAACCTG	2622
Db	2041	GAGGGGAGAGAGATGGAGAAACATTTTCACTCCGGCTGTCTCAGGGGAGGAAACCTG	2100
QY	2623	ATGCAAGTGGGCTCCCTCCCTGAGCTGTGTGTGAGCAACACTCTCTGAGGTCCCTCCAC	2682
Db	2101	ATGCAAGTGGGCTCCCTCCCTGAGCTGTGTGTGAGCAACACTCTCTGAGGTCCCTCCAC	2160

QY 2683 TGCTTGTACGAGACTCGGAACAAACGTTCTCTGACACAAAGTATGATGCCCAATTCGAAGAA 2742
Db 2161 TGCTTGTACGAGACTCGGAACAAACGTTCTCTGACACAAAGTATGATGCCCAATTCGAAGAA 2220
QY 2743 ATGGGCAATGTGTAGAAACAGACATGGAGCTCTTAGTGTGACACTTCTGCAATTAATTC 2802
Db 2221 ATGGGCAATGTGTAGAAACAGACATGGAGCTCTTAGTGTGACACTTCTGCAATTAATTC 2280
QY 2803 AGCCGACAGTGAAGAGCTTCAGCTGATTGAGGGCAGGCAGACACAGATCAACATGGAGC 2862
Db 2281 AGCCGACAGTGAAGAGCTTCAGCTGATTGAGGGCAGGCAGACACAGATCAACATGGAGC 2340
QY 2863 CCCACCATGGTATGCTCTGTTCAAGTGGGTCCCAAGTCAACAGATGCGCTATTGGCAGATTTCTC 2922
Db 2341 CCCACCATGGTATGCTCTGTTCAAGTGGGTCCCAAGTCAACAGATGCGCTATTGGCAGATTTCTC 2400
QY 2923 TTCTCGCTCTCAAGGTCAACGAAACCTGAAGAGAGCTGGAACCTAAGTGGAACTCGCTG 2982
Db 2401 TTCTCGCTCTCAAGGTCAACGAAACCTGAAGAGAGCTGGAACCTAAGTGGAACTCGCTG 2460
QY 2983 AGCCACTCTGCAAGTGAAGAGCTTTTGTAAAGACCTTGAGACGCCCTCGCTGCTCTCTGGAG 3042
Db 2461 AGCCACTCTGCAAGTGAAGAGCTTTTGTAAAGACCTTGAGACGCCCTCGCTGCTCTCTGGAG 2520
QY 3043 ACCCTGCGGTGGCTGGCTCTGAGCTTCACAGCTGAGGACTGCAAGGACCTTGCTTTGGG 3102
Db 2521 ACCCTGCGGTGGCTGGCTCTGAGCTTCACAGCTGAGGACTGCAAGGACCTTGCTTTGGG 2580
QY 3103 CTGAGAGCAACAGACCTTGACCGAGCTGAGCTTCAATGTGCTCAAGATGCT 3162
Db 2581 CTGAGAGCAACAGACCTTGACCGAGCTGAGCTTCAATGTGCTCAAGATGCT 2640
QY 3163 GGAGCCAAACCTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGGAGCTGAG 3222
Db 2641 GGAGCCAAACCTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGGAGCTGAG 2700
QY 3223 CTGGTCAGCTGTGGCTCTGAGCTGCTGCTGAGGAGCTGCGCTCTGCTTAGTGCC 3282
Db 2701 CTGGTCAGCTGTGGCTCTGAGCTGCTGCTGAGGAGCTGCGCTCTGCTTAGTGCC 2760
QY 3283 AGCCCGAGCTGAAGAGCTAGACTGCGAGAGAAACAACTGGATGACCTTGCGGTGCGA 3342
Db 2761 AGCCCGAGCTGAAGAGCTAGACTGCGAGAGAAACAACTGGATGACCTTGCGGTGCGA 2820
QY 3343 CTGCTCTGTGAGGGCTCAGGCATCTGCTGCAAACTCATAGCTGGGCTGGACAG 3402
Db 2821 CTGCTCTGTGAGGGCTCAGGCATCTGCTGCAAACTCATAGCTGGGCTGGGCTGGG 2869
QY 3403 ACAACTCTGAGTGATGAGATGAGGCGAGAACTGAGGGCTCTGAGCAGGAGAACTCTAG 3462
Db 2870 ----- 2869
QY 3463 CTGCTCATCTTACGACAGCGAAACCAAGTGTGATGACCCCTACTAGGGCTGGATAGC 3522
Db 2870 -----GGAACCAAGTGTGATGACCCCTACTAGGGCTGGATAGC 2910
QY 3523 GAGAGATGAGTAATAGCACATCTCTCACTCAAGCGCAGAGACTCGGATCAGAGGGCG 3582
Db 2911 GAGAGATGAGTAATAGCACATCTCTCACTCAAGCGCAGAGACTCGGATCAGAGGGCG 2970
QY 3583 GCTTCCATGTTGCTCAGGCTAATCTCAAACTCTGAGCTGAGCAAGATCTTCCCAATT 3642
Db 2971 GCTTCCATGTTGCTCAGGCTAATCTCAAACTCTGAGCTGAGCAAGATCTTCCCAATT 3030
QY 3643 GCTGAGATGACAGGAAAGCTCCCGAGAGTAGTACCGGTGGAACCTTTGCGTGCCT 3702
Db 3031 GCTGAGATGACAGGAAAGCTCCCGAGAGTAGTACCGGTGGAACCTTTGCGTGCCT 3090
QY 3703 TCTCTGCTCTCAAGGGACCTGCAATCAAGACCTTTGGGAGCTGACGATGATCTCTGG 3762
Db 3091 TCTCTGCTCTCAAGGGACCTGCAATCAAGACCTTTGGGAGCTGACGATGATCTCTCTG 3150

QY 3763 GGCCCCAGGGGCTGTGGCTACTGAGGTAGTTCGACAAAAGAAAGAACTTGTACGAGTT 3822
Db 3151 GGCGCTGAAGGAAATGTGGATGTTGAGTTGATTAAGAGCACAACAGATACAGCGTT 3210
QY 3823 CACTTCCCTGTAGTGGCTCTACCGTGGCCCAACACGGGTCTCTGCTTTGTGATGAGA 3882
Db 3211 TGGTTCCCACTGCTGGCTGGTATCTGTGGTCAAGCCAGCCCTCGGCTTCTGTGTAAG 3270
QY 3883 GAAGCGGTACCGTTGAGATTGAATTTCTGTGTGGGACCAAGTTCTCTGG--GTGAGATC 3939
Db 3271 GATGAGGTCAAGTACGATGCGTTTGGTTCTGGAGTCAGACCTGCGCTGGACCTG 3330
QY 3940 AACCCACAGCAGCTGATGGTGGCAGGGCTCTCTGTGGACATCAAGGCTGAGCCTGGA 3999
Db 3331 CAGCACCATGAACAGTGGTGGTGGCGGCCCTTGTGTTGATGTCACTGCGAGAGCAGAG 3390
QY 4000 GCTGTGGAAGCTGTGCACTCCCTCACCTTGTGGCTCTCCAAAGGGGCGCATGTGACACA 4059
Db 3391 GAGGCTGTGCGCGAAATCCACCTTCCCTCCCTCCAAAGGTGAGGTGAGCTC 3450
QY 4060 TCCCTGTTCCAAATGGCCCACTTTAAAGAGGAGGGGATGCTCTCTGGAGAACGCCAGG 4119
Db 3451 TCCTGTTCTCGTTGCCCATTTTAAAGATGAAGGATGGTCTCTGGAGCATCCAGCCGG 3510
QY 4120 GTGAGAGTGCATCAATAGTCTTGAAACCCAGCTTCTCCCTTGGGAGTCTCTCTG 4179
Db 3511 GTGAGAGCTTCTATGCTGTCTGAAAGCCCACTTCTCTGATGGGATCTCTCTGCTG 3570
QY 4180 AAATGATCCATATGCTGCTGGCTTCACTCCCTCACCTCTGTGGTGTGCTTTACAC 4239
Db 3571 CGGATGCGCAGTGGGACTCGGCTCTCCATCCCACTTCCCAACATGATCTAT 3630
QY 4240 CGCGTCCATCTGAGAGAACTCACCTTCCACCTCTACCTGATCCCAAGTACTGCTTCCATT 4299
Db 3631 CACCCCAACCCGAGATATTAAAGTTCCACTGTGCTTGTCCCGAGCGCTTGCTA 3690
QY 4300 CGGAAGAACTGGA 4313
Db 3691 ACAAGGCGATAGA 3704

RESULT 13
US-09-895-298-22
; Sequence 22, Application US/09895298
; Publication No. US20030078405A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 47 Human Secreted Proteins
; FILE REFERENCE: PZ035P1
; CURRENT APPLICATION NUMBER: US/09/895,298
; CURRENT FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 09/591,16
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: PCT/US99/29950
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: 60/113,006
; PRIOR FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 60/112,809
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 2657
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-895-298-22

Query Match 47.3%; Score 2572.6; DB 11; Length 2657;
Best Local Similarity 99.4%; Pred No. 0;
Matches 2614; Conservative 0; Mismatches 14; Indels 3; Gaps 3;

QY 2814 GAAGAGCTTCAAGCTGATTGAGGGCAGGACAGATCAACATGAGGCCCATGGT 2873

Db	9	GAAGAAAGCTTCAGCTGATTGAGGCGAGCGACGACAGATCAACATGGAGCCCCCACCATGGT	68
Qy	2874	AGTCCTGTTCAGSTGGGTCCAGTCCAGTACAGATGCCCTATTGGCAGATTCTTCTTCCGTCCT	2933
Db	69	AGTCCTGTTCAGSTGGGTCCAGTCCAGTCCAGTCCAGATGCCCTATTGGCAGATTCTTCTTCCGTCCT	128
Qy	2934	CAAGGTCAACAGAAACCTGAAGGAGCTGGACCTTAAGTGGAAAACTCGCTGAGCNACTCTGC	2993
Db	129	CAAGGTCAACAGAAACCTGAAGGAGCTGGACCTTAAGTGGAAAACTCGCTGAGCNACTCTGC	188
Qy	2994	AGTGAAGAGTCTTTGTAAAGCCCTGAGACGCCCTCGCTGCCTCTGTCGAGACCCCTGCGGTT	3053
Db	189	AGTGAAGAGTCTTTGTAAAGCCCTGAGACGCCCTCGCTGCCTCTGTCGAGACCCCTGCGGTT	248
Qy	3054	GGCTGGCTGTGGCTCCACAGCTGAGGACTGCAAGGACTTTGCCCTTTGGGCTGAGAGCCAA	3113
Db	249	GGCTGGCTGTGGCTCCACAGCTGAGGACTGCAAGGACTTTGCCCTTTGGGCTGAGAGCCAA	308
Qy	3114	CCAGACCCTGACCGAGCTGGACCTGAGCTTCAATGTGCTCACGATGCTCGAGCCAAACA	3173
Db	309	CCAGACCCTGACCGAGCTGGACCTGAGCTTCAATGTGCTCACGATGCTCGAGCCAAACA	368
Qy	3174	CTTTTGCAGAGACTGAGACAGCCGACTGCAAGCTACAGCGACTGCAGCTGTGTCAGCTG	3233
Db	369	CTTTTGCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCGACTGCAGCTGTGTCAGCTG	428
Qy	3234	TGGCCTCACGTCCTGACTGCTGCCAGGACCTGGCCTCTGTGCTTAGTGCCAGCCCCAGCCT	3293
Db	429	TGGCCTCACGTCCTGACTGCTGCCAGGACCTGGCCTCTGTGCTTAGTGCCAGCCCCAGCCT	488
Qy	3294	GAAGGAGCTAGACTGTCAGCAGCAAAACAACTGGATGACGTTGGCTGCGACTGCTCTGTGA	3353
Db	489	GAAGGAGCTAGACTGTCAGCAGCAAAACAACTGGATGACGTTGGCTGCGACTGCTCTGTGA	548
Qy	3354	GGGGCTCAGGATCTCTGCCTGCAAACTCATACGCTTGGGGCTGGACAGCAACTCTGAG	3413
Db	549	GGGGCTCA - GCATCTCTGCTGCAAACTCATACGCTTGGGGCTGGACAGCAACTCTGAG	607
Qy	3414	TGATGAGATGAGGACGAACTTGAGGGCCCTGGAGCAGGAGAAACCTCAGCTGTCTCATCTT	3473
Db	608	TGATGAGATGAGGACGAACTTGAGGGCCCTGGAGCAGGAGAAACCTCAGCTGTCTCATCTT	667
Qy	3474	CAGCAGACGAAACCAAGTGTGATGACCCCTACTCTGAGGGCTTGGATACGGGAGAGATGAG	3533
Db	668	CAGCAGACGAAACCAAGTGTGATGACCCCTACTGA - GGCCTGATACGGGAGAGATGAG	726
Qy	3534	TAATAGCACATCTCTCACTAAGCGGCGAGAGCTCGGATCGAGAGGGCGGCTTCCCATGT	3593
Db	727	TAATAGCACATCTCTCACTAAGCGGCGAGAGCTCGGATCGAGAGGGCGGCTTCCCATGT	786
Qy	3594	TGCTCAGGCTAACTCTAAACTCTCTGGACGTGAGCAAGATCTTCCCAATTGCTGAGATTGC	3653
Db	787	TGCTCAGGCTAACTCTAAACTCTCTGGACGTGAGCAAGATCTTCCCAATTGCTGAGATTGC	846
Qy	3654	AGAGGAAAGTCTCCCAGAGGTAGTACCGGTGGAACTCTTGTGCTGCTTCTCTGCGCTC	3713
Db	847	AGAGGAAAGTCTCCCAGAGGTAGTACCGGTGGAACTCTTGTGCTGCTTCTCTGCGCTC	906
Qy	3714	TCAAGGGGACCTGTCAATACGAAGCTTTGGGGACTGACGATGACTTCTGGGGCCCCCAGGG	3773
Db	907	TCAAGGGGACCTGTCAATACGAAGCTTTGGGGACTGACGATGACTTCTGGGGCCCCCAGGG	966
Qy	3774	GCCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTTCACATTCCTCTGT	3833
Db	967	GCCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTTCACATTCCTCTGT	1026
Qy	3834	AGCTTGGCTCTTACCGCTGGCCCCAACACGGGTCTCTGCTTTGTGATGAGAAAGCGGTGAC	3893
Db	1027	AGCTTGGCTCTTACCGCTGGCCCCAACACGGGTCTCTGCTTTGTGATGAGAAAGCGGTGAC	1086
Qy	3894	CGTTGAGATTGAATTCCTGTGTGGGACCAAGTTCCTGGGTGAGATCAACCAAGACAG	3953
Db	1087	CGTTGAGATTGAATTCCTGTGTGGGACCAAGTTCCTGGGTGAGATCAACCAAGACAG	1146

3954	CTGGATGTTGCAGGGGCTCTGCTGGACATCAAGGCTGAGCCTGGAGCTGTGGAAAGCTGT	4013
1147	CTGGATGTTGCAGGGGCTCTGCTGGACATCAAGGCTGAGCCTGGAGCTGTGGAAAGCTGT	1206
4014	GCACTCCCTCATCTTGTGCTCTCCAAGGGGCCATGTGGACACATCCCTGTTTCAAAT	4073
1207	GCACCTCCCTCACCTTGTGCTCTCCAAAGGGGGCCATGTGGACACATCCCTGTTTCAAAGT	1266
4074	GGCCCCATTTAAAGAGAGGGGATGCTCTCGGAAGAAGCACGCCAGGGTGAGCTGCATCA	4133
1267	GGCCCCATTTAAAGAGAGGGGATGCTCTCGGAAGAAGCACGCCAGGGTGAGCTGCATCA	1326
4134	CATAGTTCTGAAAACCCCAAGCTTCTCCGCCCTTGGGAGTCCTCTGAAAAATGATCCATAA	4193
1327	CATAGTTCTGAAAACCCCAAGCTTCTCCGCCCTTGGGAGTCCTCTGAAAAATGATCCATAA	1386
4194	TGCCCTGGCTTCATTCCCGTCACTCTGTGTGTGTGCTTTACCACGCGTCCATCCGTA	4253
1387	TGCCCTGGCTTCATTCCCGTCACTCTGTGTGTGTGCTTTACCACGCGTCCATCCGTA	1446
4254	GGAAGTCACCTTCCACCTCTACCTGATCCCAAGTGACTGCTCCATTCCGGAAGGAACCTGA	4313
1447	GGAAGTCACCTTCCACCTCTACCTGATCCCAAGTGACTGCTCCATTCCGGAAGGAACCTGA	1506
4314	GCTCTGCTATCGAAGCCCTGGAGAAGACAGCTGTTCTCGAGTTCATCGTTGGCCACTT	4373
1507	GCTCTGCTATCGAAGCCCTGGAGAAGACAGCTGTTCTCGAGTTCATCGTTGGCCACTT	1566
4374	GGGATCAGGATCAGGCTGCAAGTGAAGAACAAGAAAGATGAGACTCTGGTGTGGGAGGC	4433
1567	GGGATCAGGATCAGGCTGCAAGTGAAGAACAAGAAAGATGAGACTCTGGTGTGGGAGGC	1626
4434	CTTGTGAAACCGAGGATCTCATGCTCGCAACTCTGATCCCTCCAGCCCGCATPAGC	4493
1627	CTTGTGAAACCGAGGATCTCATGCTCGCAACTCTGATCCCTCCAGCCCGCATPATC	1686
4494	CGTACTCTCACCTCTGGATGCCCGCAGTTGCTGCACCTTTGTGACACAGTATCGAGACGA	4553
1687	CGTACTCTCACCTCTGGATGCCCGCAGTTGCTGCACCTTTGTGACACAGTATCGAGACGA	1746
4554	GCTGATACCGGAGTAGATCGGTGAGATTGTCTTGGACAAACTGCATGACAGAGTGCT	4613
1747	GCTGATACCGGAGTAGATCATCGGTGAGATTGTCTTGGACAAACTGCATGACAGAGTGCT	1806
4614	GAGCAGAGCAGTACAGAGGGTGTGGCTGGAACACAGGCCCCAGCAGATGCGGAA	4673
1807	GAGCAGAGCAGTACAGAGGGTGTGGCTGGAACACAGGCCCCAGCAGATGCGGAA	1866
4674	GCTGTTTCACTTGAGCCAGTCTCGGACCGGAAGTGC AAAAGATGGACTTACCAAGCCCT	4733
1867	GCTGTTTCACTTGAGCCAGTCTCGGACCGGAAGTGC AAAAGATGGACTTACCAAGCCCT	1926
4734	GAAGGAGACCCATCCTCACTTATATGAACTCTGGGAAGGGCAGCAAAAAGGAGCT	4793
1927	GAAGGAGACCCATCCTCA - CTCATTTATGAACTCTGGGAAGGGCAGCAAAAAGGAGCT	1985
4794	CTTGCCACTCAGCAGCTGAAGTATCAACACAGCCCTTGACCCCTTGAGTCTCGCTTTGG	4853
1986	CTTGCCACTCAGCAGCTGAAGTATCAACACAGCCCTTGACCCCTTGAGTCTCGCTTTGG	2045
4854	CTGACCTTCTTTGGGTCTCAGTTTTCTTCTCTGCAAAACAAGTTGCCATCTGTGTTGCCT	4913
2046	CTGACCTTCTTTGGGTCTCAGTTTTCTTCTCTGCAAAACAAGTTGCCATCTGTGTTGCCT	2105
4914	TCAGCACTAAAGTAAATGGAATCTTGATGATGCCCTTGTGGGCAATTATGTCCTATGCC	4973
2106	TCAGCACTAAAGTAAATGGAATCTTGATGATGCCCTTGTGGGCAATTATGTCCTATGCC	2165
4974	AGGGATGCCACAGGGGGCCCAGTCCAGGTGGCCTAACAGCATCTCAGGGNATGTCATC	5033
2166	AGGGATGCCACAGGGGGCCCAGTCCAGGTGGCCTAACAGCATCTCAGGGNATGTCATC	2225

QY	5034	TGAGCTGCAAGACCCCTGCGACCTCATAGAGCCTCATCTGGTGGCCACAGAGCCAA	5093	Db	30798	GTCAGAGAGAGCGTGGCCCTGATTATGTGAGAGAGAAATCAGAGACATTTAAATTGAGATCAGA	30857
Db	2226	TGAGCTGCAAGACCCCTGCGACCTCATAGAGCCTCATCTGGTGGCCACAGAGCCAA	2285	QY	1468	GACTTATTTGGCCCAAGGCTGATATCCCAAGAACTCCCATAGTATCATCTGAGGGGGCT	1527
QY	5094	GCTAGAGCCCTCCGATCCCATCCAGCGCAAGAGGAATAGGAGGACATGGAACCAT	5153	Db	30858	GACTTATTTGGCCCAAGGCTGATATCCCAAGAACTCCCATAGTATCATCTGAGGGGGCT	30917
Db	2286	AGCTTAGAGCCCTCCGATCCCATCCAGCGCAAGAGGAATAGGAGGACATGGAACCAT	2345	QY	1528	GCTGAAATTTGGGAAAGTCAACTGCGCAAGGAGGTGAAGGAAGCCCTGGGGGAGAGCCAG	1587
QY	5154	TTGCTCTTGGCTGTGTACAGGGGTGAGCCCAAAATTTGGGGTTTCAGCGTGGGAGGCCAG	5213	Db	30918	GCTGAAATTTGGGAAAGTCAACTGCGCAAGGAGGTGAAGGAAGCCCTGGGGGAGAGCCAG	30977
Db	2346	TTGCTCTTGGCTGTGTACAGGGGTGAGCCCAAAATTTGGGGTTTCAGCGTGGGAGGCCAG	2405	QY	1588	CTGTATGGGGACCGCTTCCAGCATGTCTTCTACTTTCACTTCAGTGCAGAGAGTGGCCAGTCC	1647
QY	5214	TGATTTCTGGCTTTGTGACAGGAAGATCTACAGAGCAAGCCAAACAGATTAAGTGAAG	5273	Db	30978	CTGTATGGGGACCGCTTCCAGCATGTCTTCTACTTTCACTTCAGTGCAGAGAGTGGCCAGTCC	31037
Db	2406	TGATTTCTGGCTTTGTGACAGGAAGATCTACAGAGCAAGCCAAACAGATTAAGTGAAG	2465	QY	1648	AAGGTGGTCACTGCTCGCTCAGCTCATCGGAAAGATGGGACAGCCACCTCCGCTCCCAT	1707
QY	5274	GAGTTTATTCAGAAATTAAGGAGTATCACAGCTCTTTTGAATTTGTCTAGCAGGCTT	5333	Db	31038	AAGGTGGTCACTGCTCGCTCAGCTCATCGGAAAGATGGGACAGCCACCTCCGCTCCCAT	31097
Db	2466	GAGTTTATTCAGAAATTAAGGAGTATCACAGCTCTTTTGAATTTGTCTAGCAGGCTT	2525	QY	1708	AGACAGATCTGTCTAGGCGCAGAGCGGTCTCTTCATCTCGATGGTGTAGATGAGCA	1767
QY	5334	TCAGTCTTTTACAGAAACCCCTATAAATTTAAATTTTAAATTTTAAATTTTAAATTTA	5393	Db	31098	AGACAGATCTGTCTAGGCGCAGAGCGGTCTCTTCATCTCGATGGTGTAGATGAGCA	31157
Db	2526	TCAGTCTTTTACAGAAACCCCTATAAATTTAAATTTTAAATTTTAAATTTTAAATTTA	2585	QY	1768	GGATGGGTCTTGCAGAGCGGAGTTCTGAGCTCTGCTGCACTGGAGCCAGCCACAGCGG	1827
QY	5394	AAAAATACAAAAAGAAAAATGAAAAATGAAAAATGAAAAATGAAAAATGAAAAATG	5444	Db	31158	GGATGGGTCTTGCAGAGCGGAGTTCTGAGCTCTGCTGCACTGGAGCCAGCCACAGCGG	31217
Db	2586	AAAAATACAAAAAGAAAAATGAAAAATGAAAAATGAAAAATGAAAAATGAAAAATG	2636	QY	1828	GCGGATGCACTGCTGGGCGAGTTTGTCTGGGAAAACTATATCTTCCGAGGCACTCTTCCTG	1887

RESULT 14

US-09-956-712-10
 ; Sequence 10, Application US/09956712
 ; Publication No. US20030092648A1
 ; GENERAL INFORMATION:
 ; APPLICANT: C. Frank Bennett
 ; APPLICANT: Susan M. Freier
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF NAC EXPRESSION
 ; FILE REFERENCE: RTS-0326
 ; CURRENT APPLICATION NUMBER: US/09/956,712
 ; CURRENT FILING DATE: 2001-09-19
 ; NUMBER OF SEQ ID NOS: 91
 ; SEQ ID NO 10
 ; LENGTH: 96649
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(96649)
 ; OTHER INFORMATION: n = A,T,C or G

QY	1168	TACAGAAATCAGAGAAAGAGAGAGAGAGAAATCAGAGAAAGCGAGGCCCATGGCA	1227	Db	31578	GCCAGGCTCTCCAGCTCAGCCATTTGGACCCCAAGCTCAGAGACCTCTGCTCTGGCT	31637
Db	30558	TACTTAGAAATCAGAGAAAGAGAGAGAGAGAAATCAGAGAAAGCGAGGCCCATGGCA	30617	QY	2248	GCTGAGGGCATCTGGCAAAAAAGACCCCTTTTCACTCCAGATGACCTCAGGAAGCATGGG	2307
QY	1228	GCGGTGGTAGGAACGCCCCACAGCGGCACACAGCCTACAGCCCAACCAACCCATGG	1287	Db	31638	GCTGAGGGCATCTGGCAAAAAAGACCCCTTTTCACTCCAGATGACCTCAGGAAGCATGGG	31697
Db	30618	GCGGTGGTAGGAACGCCCCACAGCGGCACACAGCCTACAGCCCAACCAACCCATGG	30677	QY	2308	TTAGATGGGGCCATCATCTCCACCTTTTGAAGTGGGTATTTCTTCAAGAGGACACCCATC	2367
QY	1288	GAGCCTCTGTGAGAGAGCCCTCTTCCACATGGCCCTGGAATATGAGATTTTAAAC	1347	Db	31698	TTAGATGGGGCCATCATCTCCACCTTTTGAAGTGGGTATTTCTTCAAGAGGACACCCATC	31757
Db	30678	GAGCCTCTGTGAGAGAGCCCTCTTCCACATGGCCCTGGAATATGAGATTTTAAAC	30737	QY	2368	CCTCTGAGCTACAGCTTTCATTCACCTCTGTTTCCAAAGAGTTCTTTTGCAGCAATGCTAT	2427
QY	1348	CAAAATTTACACAGCTGCTTCTTACAAAGACCTCACCCCAAGAGCCAAAGATCCCTG	1407	Db	31758	CCTCTGAGCTACAGCTTTCATTCACCTCTGTTTCCAAAGAGTTCTTTTGCAGCAATGCTAT	31817
Db	30738	CAAAATTTACACAGCTGCTTCTTACAAAGACCTCACCCCAAGAGCCAAAGATCCCTG	30797	QY	2428	GTCTTTGAGGATGAGAGGAGAGGTAAAATTTCTTAATTGATCATAGATTGGAAG	2487
QY	1408	GTCAGAGAGCTGGCTGATTATGTGGAGGAATCGAGGACATTTAATTGAGATCAGA	1467	Db	31818	GTCTTTGAGGATGAGAGGAGAGGTAAAATTTCTTAATTGATCATAGATTGGAAG	31877

QY 2548 GGCTCTTAAGTATGATGAGGGGAGAGAGATGAGAAATCTTTTCACTGCCGGCTGTCT 2607
 DB 31938 GGCTCTTAAGTATGATGAGGGGAGAGAGATGAGAAATCTTTTCACTGCCGGCTGTCT 31997
 QY 2608 CAGGGAGGAACTGTATGATGAGGGTCCCGTCCCTCAGCTGTGTGTCAGCCACACTCT 2667
 DB 31998 CAGGGAGGAACTGTATGATGAGGGTCCCGTCCCTCAGCTGTGTGTCAGCCACACTCT 32057
 QY 2668 CTGGAGTCCCTCACTCTGTGTACGAGACTCGGAAACAAAACGTTCTGTGACACAAAGTATG 2727
 DB 32058 CTGGAGTCCCTCACTCTGTGTACGAGACTCGGAAACAAAACGTTCTGTGACACAAAGTATG 32117
 QY 2728 GCCCATTTTCGAGAAATGGGCTGTGTAGAAACAGACATGAGCTCTTAGTGTGCACT 2787
 DB 32118 GCCCATTTTCGAGAAATGGGCTGTGTAGAAACAGACATGAGCTCTTAGTGTGCACT 32177
 QY 2788 TTCTGATTAAATTCAGCCGCCACGTTGAAGAGCTTCAGCTGATTAGGGCAGGCAGCAC 2847
 DB 32178 TTCTGATTAAATTCAGCCGCCACGTTGAAGAGCTTCAGCTGATTAGGGCAGGCAGCAC 32237
 QY 2848 AGATCAACATGAGGCCCCACCATGTTAGTCTGT 2881
 DB 32238 AGATCAACATGAGGCCCCACCATGTTAGTCTGT 32271

RESULT 15
 US-10-027-632-150437
 ; Sequence 150437, Application US/10027632
 ; Publication No. US20030204075A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; Polymorphisms in the Human Genome
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 150437
 ; LENGTH: 720
 ; TYPE: DNA
 ; ORGANISM: Human
 US-10-027-632-150437

Query Match 12.8%; Score 694.4; DB 13; Length 720;
 Best Local Similarity 99.3%; Pred. No. 4.1e-182;
 Matches 717; Conservative 1; Mismatches 2; Indels 2; Gaps 2;
 QY 1626 CTGCAGAGAGCTGGCCCGACGTCGAGTGTGAGTCTCGCTGAGCTCATCGGAAAAGATGG 1685
 DB 1 CTGCAGAGAGCTGGCCCGACGTCGAGTGTGAGTCTCGCTGAGCTCATCGGAAAAGATGG 60
 QY 1686 GACAGCACTCCGGCTCCCATTTAGACAGATCTGTCTAGCCAGAGCGGCTCTTTCAT 1745
 DB 61 GACAGCACTCCGGCTCCCATTTAGACAGATCTGTCTAGCCAGAGCGGCTCTTTCAT 120
 QY 1746 CCTCGATGGTGTAGATGAGCCAGGATGGGTCTTTCAGAGCCGAGTCTTGAGCTCTGTCT 1805

DB 121 CCTCGATGGTGTAGATGAGCCAGGATGGGTCTTCAGAGAGCCGAGTCTTGAGCTGTGTCT 180
 QY 1806 GCACTGGAGCCAGCCACAGCCGGGATGCACTGTCTGGGCGAGTTTGTCTGGGAAAATCTAT 1865
 DB 181 GCACTGGAGCCAGCCACAGCCGGGATGCACTGTCTGGGCGAGTTTGTCTGGGAAAATCTAT 240
 QY 1866 ACTTCCCGAGGCACTCTTCTGATCAGCGCTCGGACCAACAGCTCTGCAAGAACCTTCATTCC 1925
 DB 241 ACTTCCCGAGGCACTCTTCTGATCAGCGCTCGGACCAACAGCTCTGCAAGAACCTTCATTCC 300
 QY 1926 TTCTTTGAGAGGAGGCACTTGGGTAGAGTCTCTGGGTCTCTGAGTCCAGGAGAAAGA 1985
 DB 301 TTCTTTGAGAGGAGGCACTTGGGTAGAGTCTCTGGGTCTCTGAGTCCAGGAGAAAGA 360
 QY 1986 ATATTCTTACAGATATTTTACAGATGAAAGCAAGCAATTTAGAGCTTTTAGTGTGTCAA 2045
 DB 361 ATATTCTTACAGATATTTTACAGATGAAAGCAAGCAATTTAGAGCTTTTAGTGTGTCAA 420
 QY 2046 ATCAAAACAAAGAGCTCTGGGCCCTGTGTCTTGTGCTCTGGGTCTCTGGCTGGCTGCAC 2105
 DB 421 ATCAAAACAAAGAGCTCTGGGCCCTGTGTCTTGTGCTCTGGGTCTCTGGCTGGCTGCAC 480
 QY 2106 TTGCCTGATGAGCAGATGAAAGCGGAAAGAAAACCTCACAGTCTTCAAGAACACACAC 2165
 DB 481 TTGCCTGATGAGCAGATGAAAGCGGAAAGAAAACCTCACAGTCTTCAAGAACACACAC 540
 QY 2166 AACCTCTGTCTACATTTACCTTGGCCAGGCTCTCCAGCTCAGCCATTTGGGACCCAGCT 2225
 DB 541 AACCTCTGTCTACATTTACCTTGGCCAGGCTCTCCAGCTCAGCCATTTGGGACCCAGCT 600
 QY 2226 CAGAGACCTCTGCTCTCTGCTGCTGAGGGCATCTGGCAAAAAGAACCCCTTTTCAGTCC 2285
 DB 601 CAGAGACCTCTGCTCT-TGGCTGCTGAGGGCATCTGGCAAAAAGAACCC-TTTTAGTCC 658
 QY 2286 AGATGACCTCAGGAAGCATGGGTAGATGGGGCCATCATCTCCACCTTCTTGAAGATGGG 2345
 DB 659 AGATGACCTCAGGAAGCATGGGTAGATGGGGCCCTCATCTCCACCTTCTTGAAGATGGG 718
 QY 2346 TA 2347
 DB 719 TA 720

Search completed: January 29, 2004, 13:44:36
 Job time : 1744.98 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2004, 02:57:06 ; Search time 1342.51 seconds
(without alignments)
10946.442 Million cell updates/sec

Title: US-09-996-617-1
Perfect score: 5444
Sequence: 1 gccccaggcctggagaggt.....aggaataagaagttacctac 5444

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_19Jun03:*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*
- 25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5444	100.0	5523	22	AAF83651 Human CARD-7 poly
2	5444	100.0	5523	25	AB556030 cDNA encoding huma
3	5439.2	99.9	5523	25	AB555497 cDNA encoding huma
4	4879.2	89.6	5100	24	AAU47127 Pyrin domain conta
5	4765.6	87.5	5122	22	AAU12951 Human G-protein co
6	4148	76.2	4422	22	AAD02760 Human NB-ARC and C
7	4078	74.9	4194	22	AAD02761 Human NB-ARC and C
8	3952	72.6	4329	22	AAD02762 Human NB-ARC and C

ALIGNMENTS

RESULT 1

AAF83651
ID AAF83651 standard; cDNA; 5523 BP.

XX AAF83651;

AC AAF83651;

DT 23-JUL-2001 (first entry)

XX Human CARD-7 polypeptide encoding cDNA.

DE CARD-7; CARD-8; CARD-5; caspase recruitment domain; cancer; human;

XX autoimmune disorder; antiinflammatory; immunosuppressive; antiallergic;

XX antibacterial; antiviral; gene therapy; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 523..4812

FT /*tag= a

FT /product= "CARD-7"

XX WO200130813-A1.

XX PD 03-MAY-2001.

XX PF 27-OCT-2000; 2000WO-US29796.

XX PR 27-OCT-1999; 99US-0428252.

XX PA (MILL-) MILLENNIUM PHARM INC.

Human NAC beta iso
Human NAC gamma or
Human secreted pro
Human reproductiv
Human testicular a
Human pancreatic c
Human pancreatic c
Human PYRIN-8 cDNA
Pyrin domain conta
Human PYRIN-11 (PYR
Huma cDNA encoding
Human MDDT-13 enco
Human secreted pro
Human ORFX ORF1069
Human PYRIN-11 cDN
Human secreted pro
Secreted protein-e
Pyrin domain conta
Human nucleotide b
NOVX related polyn
Human secreted pro
Human PYRIN-3 cDNA
Nucleotide sequenc
Pyrin domain conta
Human foetal liver
Probe #15132 for g
Human brain expres
Human bone marrow
Probe #18558 used
Human liver single
Human genome-deriv
Human foetal liver
Probe #5882 for ge
Human brain expres
Human bone marrow
Probe #5850 for ge
Probe #5883 used t

9 3400.4 62.5 4556 22 AAD02764
10 3210.4 59.0 4466 22 AAD02765
11 2572.6 47.3 2657 21 AAA78392
12 484.4 8.9 578 22 AAL01744
13 484.4 8.9 578 23 ABL97037
14 482.8 8.9 487 24 ABV97890
15 414.4 7.6 416 24 ABV97891
16 399.8 7.3 3186 24 AAL44363
17 398 7.3 3300 24 AAL47129
18 369.8 6.8 3857 22 AAD14323
19 369.8 6.8 3857 25 ABX93556
20 357.2 6.6 2763 24 ABQ75801
21 319.6 5.9 363 21 AAC08520
22 316.2 5.8 330 21 AAC75514
23 313.4 5.8 2691 24 AAL44365
24 304.6 5.6 2847 22 AAS01487
25 304.6 5.6 2847 25 ABZ73494
26 255.2 4.7 2199 24 AAL47130
27 254 4.7 2575 24 AAI70683
28 254 4.7 2767 25 ABT16018
29 254 4.7 3172 24 ABN99366
30 252.4 4.6 3368 24 AAL44366
31 245.8 4.5 2614 25 ABV72512
32 241.8 4.4 1557 24 AAL47143
33 237 4.4 237 22 ABA69780
34 237 4.4 237 22 AAK17979
35 237 4.4 237 22 AAK43859
36 237 4.4 237 22 AAI49872
37 237 4.4 237 23 ABS43511
38 237 4.4 237 24 ABS18088
39 237 4.4 237 22 ABA58276
40 226.4 4.2 469 22 ABA27416
41 226.4 4.2 469 22 AAK06371
42 226.4 4.2 469 22 AAK32039
43 226.4 4.2 469 22 AAI15917
44 226.4 4.2 469 22 AAI37897
45 226.4 4.2 469 22 AAI37897

PI Bertin J;
 XX WPI: 2001-343340/36.
 DR P-PSDB; AAB62571.
 DR
 XX
 PT Identifying a modulator of interaction between caspase recruitment
 PT domain (CARD)-7 and CARD-5, for treating autoimmune disorders,
 PT comprises measuring the binding of CARD-7 and CARD-5 in the presence of
 PT the compound -
 PT
 XX
 PS Disclosure; Fig 1A-D; 80pp; English.
 XX
 CC The invention relates to identifying compounds, that modulate interaction
 CC between caspase recruitment domain (CARD)-7 and CARD-5. The method
 CC involves measuring the binding of CARD-7 and CARD-5 in the presence of
 CC the compound (an increase in the binding of CARD-7 to CARD-5 in the
 CC presence of the compound compared to the binding in the absence of the
 CC compound indicates that the compound is a modulator of CARD-7-CARD-5
 CC interaction). Modulators of CARD-7 and CARD-8 expression or activity can
 CC be used to treat or diagnose disorders such as cancers, bacterial or
 CC viral infections, autoimmune disorders (systemic lupus erythematosus,
 CC immune-mediated glomerulonephritis or arthritis), inflammatory disorders,
 CC organ-specific autoimmunity including multiple sclerosis, Hashimoto's
 CC thyroiditis, or Grave's disease, psoriasis, graft rejection, allergies.
 CC CARD-7 and CARD-8 are useful as modulating agents in regulating a variety
 CC of cellular processes including cell growth and cell death. The present
 CC sequence represents a cDNA encoding the human CARD-7 polypeptide.
 XX
 SQ Sequence 5523 BP; 1350 A; 1530 C; 1481 G; 1162 T; 0 other;
 Query Match 100.0%; Score 5444; DB 22; Length 5523;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 5444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCCAGGCGCTGGAGAGTCTGAAGAAACCTGGGAGCCAGAGCGCGGGCTCCACTCT 60
 DB 1 GCGCCAGGCGCTGGAGAGTCTGAAGAAACCTGGGAGCCAGAGCGCGGGCTCCACTCT 60

QY 61 GGGTTCTGAAGGCCAATTCCTGCTCTGGGCTCTCCACCCACCTCTTCTCAGCCTT 120
 DB 61 GGGTTCTGAAGGCCAATTCCTGCTCTGGGCTCTCCACCCACCTCTTCTCAGCCTT 120

QY 121 GCAGCTCAAGGGTTGATCTCAGAGTCCAGGAGCCAGAGAGGGAAGAAATCTGAGGAACA 180
 DB 121 GCAGCTCAAGGGTTGATCTCAGAGTCCAGGAGCCAGAGAGGGAAGAAATCTGAGGAACA 180

QY 181 CAGAACAGTGAGGGTGGCCACACCCCATCTCCCGTCAACATCTCCCTCACTC 240
 DB 181 CAGAACAGTGAGGGTGGCCACACCCCATCTCCCGTCAACATCTCCCTCACTC 240

QY 241 CCTCCCTGCTGGCCCTGGAGCCCATCCAGAGCTCCCTATCAGCTGACTTCTTCCAGT 300
 DB 241 CCTCCCTGCTGGCCCTGGAGCCCATCCAGAGCTCCCTATCAGCTGACTTCTTCCAGT 300

QY 301 GTCTTGCAGGCCCTCTGGGCTCTCCCTCCCTGGCTTTTCTACCACTCCCTCTAT 360
 DB 301 GTCTTGCAGGCCCTCTGGGCTCTCCCTCCCTGGCTTTTCTACCACTCCCTCTAT 360

QY 361 CGGCGTCTATCTGTAGGTCCCTGGGATTTAATAAATCTGGTTCGGAATGCTGAATAAGA 420
 DB 361 CGGCGTCTATCTGTAGGTCCCTGGGATTTAATAAATCTGGTTCGGAATGCTGAATAAGA 420

QY 421 GACGGTAAAGCCAGGAAAGGACAGCACTGTTCTGCTGCTGCTGATACCTCACCAC 480
 DB 421 GACGGTAAAGCCAGGAAAGGACAGCACTGTTCTGCTGCTGCTGATACCTCACCAC 480

QY 481 CTGGGAAACATCCCCAGACACCTCTTAATCTCCGGACAGAGATGGCTGGCGAGCTGG 540
 DB 481 CTGGGAAACATCCCCAGACACCTCTTAATCTCCGGACAGAGATGGCTGGCGAGCTGG 540

QY 541 GCGCGCTGGCTGTACTTGGAGTTCTGAGTTCTGAGAGAGGAGGAGTGAAGAGGTTCCAGCTT 600
 DB 541 GCGCGCTGGCTGTACTTGGAGTTCTGAGTTCTGAGAGAGGAGGAGTGAAGAGGTTCCAGCTT 600

QY 601 CTGCTCGCAATAAAGCGCACTCCAGAGAGCTCTTCGGGTGAGACACCCGCTCAGCCAGAG 660
 DB 601 CTGCTCGCAATAAAGCGCACTCCAGAGAGCTCTTCGGGTGAGACACCCGCTCAGCCAGAG 660

QY 661 AAGACAGTGTGATGAGAGTGGCTCTGCTACCTGCTGCTCAGTATGAGGAGAGCGGCGC 720
 DB 661 AAGACAGTGTGATGAGAGTGGCTCTGCTACCTGCTGCTCAGTATGAGGAGAGCGGCGC 720

QY 721 TGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGCTGAGGTCACTGCTGCGCCCAAGCC 780
 DB 721 TGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGCTGAGGTCACTGCTGCGCCCAAGCC 780

QY 781 CAGGAAGGGGCGAGGCGCACTCTCCCTCATTTCCCTTACAGCCCAAGTGAACCCACCTGGG 840
 DB 781 CAGGAAGGGGCGAGGCGCACTCTCCCTCATTTCCCTTACAGCCCAAGTGAACCCACCTGGG 840

QY 841 TCTCCAGCCAAACCCACTCCACCGCAGTGTAAATGCGCTCGATCCATGAAATTGCGCGC 900
 DB 841 TCTCCAGCCAAACCCACTCCACCGCAGTGTAAATGCGCTCGATCCATGAAATTGCGCGC 900

QY 901 GGGTGCAACCCAGGGCTCAGAGAGAGGGTTTGGAGACAGCTGCTGACACATCTGGAGCG 960
 DB 901 GGGTGCAACCCAGGGCTCAGAGAGAGGGTTTGGAGACAGCTGCTGACACATCTGGAGCG 960

QY 961 CGCTGGAGAGAAATCTCTGCTCTACTCTCTTACCAAGCTCTTCCAAGCTCCCGCAGACCAT 1020
 DB 961 CGCTGGAGAGAAATCTCTGCTCTACTCTCTTACCAAGCTCTTCCAAGCTCCCGCAGACCAT 1020

QY 1021 GAGTCTCCAAGCCAGGAGTCAACCAAGCCGCCACATCCACAGCAGTCTGCGGAGCTGG 1080
 DB 1021 GAGTCTCCAAGCCAGGAGTCAACCAAGCCGCCACATCCACAGCAGTCTGCGGAGCTGG 1080

QY 1081 GGATCCCACTCAGCCCGCAGCTTAGCACCAGAGAGAGGAGGCTCTGGGAGCCCAATGG 1140
 DB 1081 GGATCCCACTCAGCCCGCAGCTTAGCACCAGAGAGAGGAGGCTCTGGGAGCCCAATGG 1140

QY 1141 CCTCTGGATGAAACGTCTAGGAAATTTTACTACACAGAAATCAGAGAAAGAGAGAGAGAAA 1200
 DB 1141 CCTCTGGATGAAACGTCTAGGAAATTTTACTACACAGAAATCAGAGAAAGAGAGAGAGAAA 1200

QY 1201 TCAGAGAAAGCAGGCGCCCACTGGGAGCGGTGGTAGGAAACGCCCCCAAGCGCACACC 1260
 DB 1201 TCAGAGAAAGCAGGCGCCCACTGGGAGCGGTGGTAGGAAACGCCCCCAAGCGCACACC 1260

QY 1261 AGCCTACAGCCCAACCAACCCATGGGAGCTTCTGTGAGAGAGAGCTCTGTTCACA 1320
 DB 1261 AGCCTACAGCCCAACCAACCCATGGGAGCTTCTGTGAGAGAGAGCTCTGTTCACA 1320

QY 1321 TGGCCCTGGAATAATGAGGATTTTAAACCAAAATTTTCAACAGCTGCTACTTCTACAAAGA 1380
 DB 1321 TGGCCCTGGAATAATGAGGATTTTAAACCAAAATTTTCAACAGCTGCTACTTCTACAAAGA 1380

QY 1381 CCTCAACCCAGAGCAAGATCCCTGCTCAAGAGAGCTGGCCTGATTATGTGAGGAG 1440
 DB 1381 CCTCAACCCAGAGCAAGATCCCTGCTCAAGAGAGCTGGCCTGATTATGTGAGGAG 1440

QY 1441 AATCGAGGACATTTAAATTGAGATCAGAGACTTTTGGCCCGAGGCTGGATACCCAGAA 1500
 DB 1441 AATCGAGGACATTTAAATTGAGATCAGAGACTTTTGGCCCGAGGCTGGATACCCAGAA 1500

QY 1501 CCTCGCATAGTCATCTGAGGGGCTGCTGGAATTTGGGAAGTCAACACTGGCCAGGCG 1560
 DB 1501 CCTCGCATAGTCATCTGAGGGGCTGCTGGAATTTGGGAAGTCAACACTGGCCAGGCG 1560

QY 1561 GTGAAGGAAGCTCTGGGAGAGGCGCAGCTGTATGGGAGCCGCTTCCAGCATGTCTTCTAC 1620
 DB 1561 GTGAAGGAAGCTCTGGGAGAGGCGCAGCTGTATGGGAGCCGCTTCCAGCATGTCTTCTAC 1620

QY 1621 TTCAGCTCAGAGAGCTGGCCAGTCCAGAGTGTGAGTCTCGCTGAGCTCATCGGAAAA 1680
 DB 1621 TTCAGCTCAGAGAGCTGGCCAGTCCAGAGTGTGAGTCTCGCTGAGCTCATCGGAAAA 1680

QY 1681 GATGGAGAGCACTCCGGCTCCCATATTAGACAGATCCTGTCTAGGCGAGCGGCTGTCTC 1740
 Db 1681 GATGGAGAGCACTCCGGCTCCCATATTAGACAGATCCTGTCTAGGCGAGCGGCTGTCTC 1740
 QY 1741 TTCACTCCTCGATGGGTAGATGAGCCAGGATGGGTCTTTCAGGAGCCGAGTTCTCAGCTC 1800
 Db 1741 TTCACTCCTCGATGGGTAGATGAGCCAGGATGGGTCTTTCAGGAGCCGAGTTCTCAGCTC 1800
 QY 1801 TGTCTGCACTGGAGCCAGCCACAGCGCGGCGATGCACTGCTGGCGAGTTTCTGGGGAAA 1860
 Db 1801 TGTCTGCACTGGAGCCAGCCACAGCGCGGCGATGCACTGCTGGCGAGTTTCTGGGGAAA 1860
 QY 1861 ACTATATCTCCGAGGCACTCCTCTGTATCAAGCTCGGACCAAGCTCTGAGAACCTC 1920
 Db 1861 ACTATATCTCCGAGGCACTCCTCTGTATCAAGCTCGGACCAAGCTCTGAGAACCTC 1920
 QY 1921 ATTCTCTTTTGGAGCAGCAGCTTGGGTAGAGTCTTGGGCTCTCTGAGTCCAGCAGG 1980
 Db 1921 ATTCTCTTTTGGAGCAGCAGCTTGGGTAGAGTCTTGGGCTCTCTGAGTCCAGCAGG 1980
 QY 1981 AAGGAATATTTCTACAGATATTTTACAGATGAAGGCAAGCAATTAGAGCCTTTAGGTTG 2040
 Db 1981 AAGGAATATTTCTACAGATATTTTACAGATGAAGGCAAGCAATTAGAGCCTTTAGGTTG 2040
 QY 2041 GTCAATCAAAAGAGCTCTGGGCCCTGTGTCTTGTGCCCCTGGGTGTCCTGGCTGGCC 2100
 Db 2041 GTCAATCAAAAGAGCTCTGGGCCCTGTGTCTTGTGCCCCTGGGTGTCCTGGCTGGCC 2100
 QY 2101 TGCATTTGCTGATGACAGCAGATGAAGCGAGGAAAGCACTCACACTGACTTCCAGACC 2160
 Db 2101 TGCATTTGCTGATGACAGCAGATGAAGCGAGGAAAGCACTCACACTGACTTCCAGACC 2160
 QY 2161 ACCAAGCCTCTGTCTACATTAATCTTGGCCAGGCTCTCAAGCTCAGCCATTGGAGACC 2220
 Db 2161 ACCAAGCCTCTGTCTACATTAATCTTGGCCAGGCTCTCAAGCTCAGCCATTGGAGACC 2220
 QY 2221 CAGCTCAGAGACCTTGCTCTCTGGCTGCTGAGGCACTCGGCAAAAGAACCTTTTC 2280
 Db 2221 CAGCTCAGAGACCTTGCTCTCTGGCTGCTGAGGCACTCGGCAAAAGAACCTTTTC 2280
 QY 2281 AGTCAGATGACCTCAGGAAGCATGGGTAGATGGGGCCATCATCTCCACCTTCTTGAG 2340
 Db 2281 AGTCAGATGACCTCAGGAAGCATGGGTAGATGGGGCCATCATCTCCACCTTCTTGAG 2340
 QY 2341 ATGGGTATTCTTCAAGAGCAACCCATCCTCTGAGCTACAGCTTCACTTCACTCTGTTTC 2400
 Db 2341 ATGGGTATTCTTCAAGAGCAACCCATCCTCTGAGCTACAGCTTCACTTCACTCTGTTTC 2400
 QY 2401 CAAGAGTCTTTTGCAGCAATGTCTATGTCTTGGAGGATGAGAAGGGGAGAGGTAAACAT 2460
 Db 2401 CAAGAGTCTTTTGCAGCAATGTCTATGTCTTGGAGGATGAGAAGGGGAGAGGTAAACAT 2460
 QY 2461 TCTAATTTGATCATAGATTGGAAAGACGCTAGAGCATATGGAATACATGSCCTGTTT 2520
 Db 2461 TCTAATTTGATCATAGATTGGAAAGACGCTAGAGCATATGGAATACATGSCCTGTTT 2520
 QY 2521 GGGGCATCAACACACGTTTCTTATTTGGGCTCTTAAAGTATGAGGGGAGAGAGATG 2580
 Db 2521 GGGGCATCAACACACGTTTCTTATTTGGGCTCTTAAAGTATGAGGGGAGAGAGATG 2580
 QY 2581 GAGAACATCTTTTCACTCGCGGCTGTCTCAGGGGAGGAAACCTGATGAGTGGGTCCCGTCC 2640
 Db 2581 GAGAACATCTTTTCACTCGCGGCTGTCTCAGGGGAGGAAACCTGATGAGTGGGTCCCGTCC 2640
 QY 2641 CTGACGTGCTGTGAGCCACACTCTCTGAGTCCCTCACTGCTTGTACGAGACTCGG 2700
 Db 2641 CTGACGTGCTGTGAGCCACACTCTCTGAGTCCCTCACTGCTTGTACGAGACTCGG 2700
 QY 2701 AACAAAACGTTCTCTGACAGAGTATGGCCCAATTCGAGAAATGGCATGTGTAGAA 2760
 Db 2701 AACAAAACGTTCTCTGACAGAGTATGGCCCAATTCGAGAAATGGCATGTGTAGAA 2760
 QY 2761 ACAGACATGGAGCTCTTAGTGTGCACTTTCTGCAATTAATTTAGCCGCCACCGTGAAGAG 2820

Db 2761 ACAGACATGGAGCTCTTAGTGTGCACTTTCTGCAATTAATTTAGCCGCCACCGTGAAGAG 2820
 QY 2821 CTTCACTCTGATTTAGGGCAGGACAGATCAACATGGAGCCCAACCATGTTAGTCTCTG 2880
 Db 2821 CTTCACTCTGATTTAGGGCAGGACAGATCAACATGGAGCCCAACCATGTTAGTCTCTG 2880
 QY 2881 TTCAGGTGGTCCAGTACAGATGCTTATTTGGAGATTTCTTCTCCGTCTCAAGTCT 2940
 Db 2881 TTCAGGTGGTCCAGTACAGATGCTTATTTGGAGATTTCTTCTCCGTCTCAAGTCT 2940
 QY 2941 ACCAGAACCTGAAGGAGCTGGA CTTAAGTGGAACTCGCTGAGCCACTCTGCAAGTGAAG 3000
 Db 2941 ACCAGAACCTGAAGGAGCTGGA CTTAAGTGGAACTCGCTGAGCCACTCTGCAAGTGAAG 3000
 QY 3001 AGTCTTTTGAAGACCTGAGACGCCCTCGCTGAGACCCCTCGGTTGGCTGGC 3060
 Db 3001 AGTCTTTTGAAGACCTGAGACGCCCTCGCTGAGACCCCTCGGTTGGCTGGC 3060
 QY 3061 TGTGGCTCTCAAGCTGAGGACTGCAAGGACCTTGGCTTTGGCTGAGAGCCAAACAGACC 3120
 Db 3061 TGTGGCTCTCAAGCTGAGGACTGCAAGGACCTTGGCTTTGGCTGAGAGCCAAACAGACC 3120
 QY 3121 CTGACCGAGCTGGA CTTGAGCTTCAATGTCTCAGGATGCTGGAGCCAAACACTTTTC 3180
 Db 3121 CTGACCGAGCTGGA CTTGAGCTTCAATGTCTCAGGATGCTGGAGCCAAACACTTTTC 3180
 QY 3181 CAGAGACTGAGACAGCCGAGCTGCAAGCTACAGGCACTGCAAGCTGCTGAGTGGCTC 3240
 Db 3181 CAGAGACTGAGACAGCCGAGCTGCAAGCTACAGGCACTGCAAGCTGCTGAGTGGCTC 3240
 QY 3241 ACCTCTGCTGCTGAGGACCTGCTGCTTGTGCTTGTGCTGAGCCAGCCAGCTGAAGAG 3300
 Db 3241 ACCTCTGCTGCTGAGGACCTGCTGCTTGTGCTTGTGCTGAGCCAGCCAGCTGAAGAG 3300
 QY 3301 CTAGACCTGAGAGCAACAACTTGATGACGTTGGCTGCGACTGCTCTGTGAGGGCTC 3360
 Db 3301 CTAGACCTGAGAGCAACAACTTGATGACGTTGGCTGCGACTGCTCTGTGAGGGCTC 3360
 QY 3361 AGGCATCTGCTGCTGCAAACTCATACGCTGGGGCTGGACAGCAAACTCTGAGTATGAG 3420
 Db 3361 AGGCATCTGCTGCTGCAAACTCATACGCTGGGGCTGGACAGCAAACTCTGAGTATGAG 3420
 QY 3421 ATGAGCAGGAACTGAGGGCCCTTGAGCAGGAGAAACCTCAGCTGCTCATCTTACAGAGA 3480
 Db 3421 ATGAGCAGGAACTGAGGGCCCTTGAGCAGGAGAAACCTCAGCTGCTCATCTTACAGAGA 3480
 QY 3481 CGGAAACCAAGTGTATGACCCCTTACTGAGGSCCTGGATACGGGAGATGAGTAAATAGC 3540
 Db 3481 CGGAAACCAAGTGTATGACCCCTTACTGAGGSCCTGGATACGGGAGATGAGTAAATAGC 3540
 QY 3541 ACATCTCACTCAAGCGGACAGACTCGGATCAGAGAGGGGGCTTCCCATGTTCTCAG 3600
 Db 3541 ACATCTCACTCAAGCGGACAGACTCGGATCAGAGAGGGGGCTTCCCATGTTCTCAG 3600
 QY 3601 GCTAATCTCAAACTCTGGAAGTCTTCCCAATTTGCTGAGATTGAGAGGAA 3660
 Db 3601 GCTAATCTCAAACTCTGGAAGTCTTCCCAATTTGCTGAGATTGAGAGGAA 3660
 QY 3661 AGCTCCAGAGGTAGTACCGGTGAACTCTTGTGCTGCTTCTCTGCTCTCAAGGG 3720
 Db 3661 AGCTCCAGAGGTAGTACCGGTGAACTCTTGTGCTGCTTCTCTGCTCTCAAGGG 3720
 QY 3721 GACCTGCAATCGAAGCTTTGGGGATGACGATGACTTCTGGGGCCCAACCGGGCTGTG 3780
 Db 3721 GACCTGCAATCGAAGCTTTGGGGATGACGATGACTTCTGGGGCCCAACCGGGCTGTG 3780
 QY 3781 GCTACTGAGGTAGTTGACAAAAGAAAGAACTTGTACCGAGTTCACTTCCCTGAGTGGC 3840
 Db 3781 GCTACTGAGGTAGTTGACAAAAGAAAGAACTTGTACCGAGTTCACTTCCCTGAGTGGC 3840
 QY 3841 TCCTACCGCTGCCCAACCGGCTCTCTGCTTGTGATGAGAGAGCGGTGACCGTTGAG 3900

Db 3841 TCCTACCGCTGGCCCAACACGGGTCTCTGCTTTGTGATGAGAAAGCGGTGACCGTTGAG 3900
 QY 3901 ATTGAATTCCTGTGTGGGACCAAGTTCTCTGGGTGAGATCAACCCACAGACAGCTGGATG 3960
 Db 3901 ATTGAATTCCTGTGTGGGACCAAGTTCTCTGGGTGAGATCAACCCACAGACAGCTGGATG 3960
 QY 3961 GTGGCAGGSCCTCTGCTGGACATCAAGGCTGAGCCTGGAGCTGTGGAAGCTGTGACCTTC 4020
 Db 3961 GTGGCAGGSCCTCTGCTGGACATCAAGGCTGAGCCTGGAGCTGTGGAAGCTGTGACCTTC 4020
 QY 4021 CCTCACTTTGTGGCTCTCAAGGGGCGCATGTGGACACATCCCTGTTCCTCAATATGGCCAC 4080
 Db 4021 CCTCACTTTGTGGCTCTCAAGGGGCGCATGTGGACACATCCCTGTTCCTCAATATGGCCAC 4080
 QY 4081 TTTAAAGAGGAGGGATGCTCTCGGAGAGCCAGCCAGGCTGGAGCTGCATCACATAGTT 4140
 Db 4081 TTTAAAGAGGAGGGATGCTCTCGGAGAGCCAGCCAGGCTGGAGCTGCATCACATAGTT 4140
 QY 4141 CTGGAAAACCCAGCTTCTCCCTTGGGAGTCTCTCTGAAAATGATCCATATATGCCCTG 4200
 Db 4141 CTGGAAAACCCAGCTTCTCCCTTGGGAGTCTCTCTGAAAATGATCCATATATGCCCTG 4200
 QY 4201 CGCTTCATTCCTGCTCACTCTGTGTGTGCTTTACCAACCGGTTCATCTGAGGAAGTC 4260
 Db 4201 CGCTTCATTCCTGCTCACTCTGTGTGTGCTTTACCAACCGGTTCATCTGAGGAAGTC 4260
 QY 4261 ACCTTCACCTCTACCTGATCCCAAGTGTCTCAATTCGGAAGGAGTGGAGCTCTGC 4320
 Db 4261 ACCTTCACCTCTACCTGATCCCAAGTGTCTCAATTCGGAAGGAGTGGAGCTCTGC 4320
 QY 4321 TATCGAAGCCCTGGAGAACACACAGCTGTCTCGGAGTCTCACTTGGCCACCTTGGGATCA 4380
 Db 4321 TATCGAAGCCCTGGAGAACACACAGCTGTCTCGGAGTCTCACTTGGCCACCTTGGGATCA 4380
 QY 4381 GGGATCAGCTGCAAGTGAAGAACAAAGATGAGACTCTGGTGTGGAGCCCTTGGTG 4440
 Db 4381 GGGATCAGCTGCAAGTGAAGAACAAAGATGAGACTCTGGTGTGGAGCCCTTGGTG 4440
 QY 4441 AAACGAGGAGTCTCATGCTGCGACTCTCTGATCCCTCCAGCCCGCATGCGTACT 4500
 Db 4441 AAACGAGGAGTCTCATGCTGCGACTCTCTGATCCCTCCAGCCCGCATGCGTACT 4500
 QY 4501 TCACCTCTGGATGCCCGCAGTGTCTGCACTTTGTGGACAGTATCGAGCAGCTGATA 4560
 Db 4501 TCACCTCTGGATGCCCGCAGTGTCTGCACTTTGTGGACAGTATCGAGCAGCTGATA 4560
 QY 4561 GCCGAGTGACATCGGTGGAGTGTCTTGGACAACTGATGAGAGTGTGAGAGCTGAGCCAG 4620
 Db 4561 GCCGAGTGACATCGGTGGAGTGTCTTGGACAACTGATGAGAGTGTGAGAGCTGAGCCAG 4620
 QY 4621 GAGCAGTACGAGGCTGTGCTGAGAACACAGAGCCCGCAGATGCGGAGCTGTTTC 4680
 Db 4621 GAGCAGTACGAGGCTGTGCTGAGAACACAGAGCCCGCAGATGCGGAGCTGTTTC 4680
 QY 4681 AGCTTGAGCCAGTCTGCGGACCGAGTGTCAAGATGAGTCTTACCAAGCTGTAAGGAG 4740
 Db 4681 AGCTTGAGCCAGTCTGCGGACCGAGTGTCAAGATGAGTCTTACCAAGCTGTAAGGAG 4740
 QY 4741 ACCCATCTCTCACTCATATTGGAACCTCTGGGAGAAAGGGGAGGAGTCTCTGCA 4800
 Db 4741 ACCCATCTCTCACTCATATTGGAACCTCTGGGAGAAAGGGGAGGAGTCTCTGCA 4800
 QY 4801 CTCAGCAGCTGAGATCAACACCGCCCTTGACCCCTGAGTCTGCTGGCTTTGGCTGACCC 4860
 Db 4801 CTCAGCAGCTGAGATCAACACCGCCCTTGACCCCTGAGTCTGCTGGCTTTGGCTGACCC 4860
 QY 4861 TTTCTTTGGGTCTCAGTTCTCTCTGCAAAACAAAGTTGCACTGCTGTTTGGCTTCCAGCA 4920
 Db 4861 TTTCTTTGGGTCTCAGTTCTCTCTGCAAAACAAAGTTGCACTGCTGTTTGGCTTCCAGCA 4920
 QY 4921 CTAAGATTAAGGAATTCATGATGCTTTGCTGGGCAATATGTGTCATGCCAGGATG 4980
 Db 4921 CTAAGATTAAGGAATTCATGATGCTTTGCTGGGCAATATGTGTCATGCCAGGATG 4980

RESULT 2
 ABS56030

ID ABS56030 standard; cDNA; 5523 BP.

XX AC ABS56030;

XX DT 10-JAN-2003 (first entry)

XX CDNA encoding human caspase recruitment domain-7 (CARD-7).

XX Human; caspase activity; caspase recruitment domain-7; CARD-7;
 KW caspase-1; pseudo-interleukin-1 beta converting enzyme; IL-1beta;
 KW pseudotICE; ICEBERG; cell growth; cell death; inflammation;
 KW apoptosis; caspase activation; cancer; follicular lymphoma;
 KW leukaemia; melanoma; colon cancer; lung carcinoma; viral infection;
 KW autoimmune disease; systemic lupus erythematosus; reactive arthritis;
 KW human immunodeficiency virus infection; HIV infection; AIDS;
 KW Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
 KW myocardial infarction; stroke; inflammatory disorder; Crohn's disease;
 KW insulin dependent diabetes mellitus; multiple sclerosis; psoriasis;
 KW graft rejection; allergic rhinitis; food allergy; conjunctivitis;
 KW glomerular nephritis; cytostatic; virucide; immunosuppressive;
 KW dermatological; nephrotropic; neuroprotective; cardiact; gene; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 523..4812

XX FT /*tag= a

XX FT /product= "CARD-7"

XX PN US2002128198-A1.

XX PD 12-SEP-2002.

XX PF 27-NOV-2001; 2001US-0996617.

XX PR 28-JUN-1999; 99US-0340620.

PR 27-OCT-1999; 99US-0428252.

1381 CCTCACCCAGAGCCAAAGATCCCTGGTCAAGAGAAAGCTGGCCCTGATTATGTGGAGGAG 1440
1441 AATCGAGGACATTTAATTGAGATCAGAGACTTATTTGGCCAGGCCCTGATACCCAGAA 1500
1441 AATCGAGGACATTTAATTGAGATCAGAGACTTATTTGGCCAGGCCCTGATACCCAGAA 1500
1501 CCTCGCATAGTCAATATGACAGGGGGTGTCTGGAATTTGGGAATTCACACCTGGCCAGGCGAG 1560
1501 CCTCGCATAGTCAATATGACAGGGGGTGTCTGGAATTTGGGAATTCACACCTGGCCAGGCGAG 1560
1561 GTGAAGGAAGCCCTGGGGGAGAGCCAGCTGTATGGGGACCGCTTCAGCATGTCTTCTAC 1620
1561 GTGAAGGAAGCCCTGGGGGAGAGCCAGCTGTATGGGGACCGCTTCAGCATGTCTTCTAC 1620
1621 TTCAGCTGACAGAGAGCTGGCCCAAGTCCAAAGTGGTGAAGTCTCGCTAGCTCATCGGAATA 1680
1621 TTCAGCTGACAGAGAGCTGGCCCAAGTCCAAAGTGGTGAAGTCTCGCTAGCTCATCGGAATA 1680
1681 GATGGACAGCCACTCCGGCTCCCATTTAGACAGATCCTGTCTAGGCCAGAGCGGCTGCTC 1740
1681 GATGGACAGCCACTCCGGCTCCCATTTAGACAGATCCTGTCTAGGCCAGAGCGGCTGCTC 1740
1741 TTCATCCTCGATGGTGTAGATGAGCCAGGATGGGTCTTCAGGAGCCGAGTTCTGAGCTC 1800
1741 TTCATCCTCGATGGTGTAGATGAGCCAGGATGGGTCTTCAGGAGCCGAGTTCTGAGCTC 1800
1801 TGTCTGCACCTGAGCCAGCCACAGCCGGGGATGCACTGCTGGCGAGTTGCTGGGAAA 1860
1801 TGTCTGCACCTGAGCCAGCCACAGCCGGGGATGCACTGCTGGCGAGTTGCTGGGAAA 1860
1861 ACTATACTTCCCGAGGCATCTTCTGATCAGGCTCGGACACAGCTCTGACAGAACCTC 1920
1861 ACTATACTTCCCGAGGCATCTTCTGATCAGGCTCGGACACAGCTCTGACAGAACCTC 1920
1921 ATTCTCTTTTGGAGCAGGCACGTTGGGTAGAGGTCTCTGGGTTCCTGAGTCCAGCAGG 1980
1921 ATTCTCTTTTGGAGCAGGCACGTTGGGTAGAGGTCTCTGGGTTCCTGAGTCCAGCAGG 1980
1981 AAGGAATATTTTACAGATATTTACAGATGAAGGCAAGCAATTAGAGCCTTTAGTTG 2040
1981 AAGGAATATTTTACAGATATTTACAGATGAAGGCAAGCAATTAGAGCCTTTAGTTG 2040
2041 GTCAATCAAAAGAGCTCTGGGCCCTGTGTCTTGTGCTCGCTGGGTGCTCGCTGGCC 2100
2041 GTCAATCAAAAGAGCTCTGGGCCCTGTGTCTTGTGCTCGCTGGGTGCTCGCTGGCC 2100
2101 TGCACCTTGCCTGATGACAGAGATGAAGCGGAAGGAAAACTCACACTGACTTCCAGACC 2160
2101 TGCACCTTGCCTGATGACAGAGATGAAGCGGAAGGAAAACTCACACTGACTTCCAGACC 2160
2161 ACCAAGCCCTGTCTACATTAACCTTGGCCAGGCTCTCAAGCTCAGGCATTTGGAGCC 2220
2161 ACCAAGCCCTGTCTACATTAACCTTGGCCAGGCTCTCAAGCTCAGGCATTTGGAGCC 2220
2221 CAGCTCAGAGACCTCTGCTCTCTGGCTGTGAGGGCATCTGCAAAAAAGACCTTTTC 2280
2221 CAGCTCAGAGACCTCTGCTCTCTGGCTGTGAGGGCATCTGCAAAAAAGACCTTTTC 2280
2281 AGTCCAGATGACCTCAGGAAGCATGGGTTAGATGGGGCCATCATCTCCACCTTTTGAAG 2340
2281 AGTCCAGATGACCTCAGGAAGCATGGGTTAGATGGGGCCATCATCTCCACCTTTTGAAG 2340
2341 ATGGGTATTTTCAAGAGACCCCATCCCTCTGAGCTACAGCTTATTCACCTCTGTTTC 2400
2341 ATGGGTATTTTCAAGAGACCCCATCCCTCTGAGCTACAGCTTATTCACCTCTGTTTC 2400
2401 CAAGAGTCTTTGACAGCAATGCTTATGCTTGGAGGATGAGAGGGGAGAGGTAAACAT 2460
2401 CAAGAGTCTTTGACAGCAATGCTTATGCTTGGAGGATGAGAGGGGAGAGGTAAACAT 2460
2461 TCTAATTGCATATAGATTTGGAAGAGAGCGCTAGAGCATATGGAATACATGGCCCTGTTT 2520

2461 TCTAATTGCATATAGATTTGGAAGAGAGCGCTAGAGCATATGGAATACATGGCCCTGTTT 2520
2521 GGGGATCAAAACACAGCTTTCTCTATTGGCCCTGTTTAACTGATGAGGGGAGAGAGATG 2580
2521 GGGGATCAAAACACAGCTTTCTCTATTGGCCCTGTTTAACTGATGAGGGGAGAGAGATG 2580
2581 GAGAAACATCTTTTCACTGCGGCTGTCTCAGGGGAGAACTCTGATGACGTGGTCCCGTCC 2640
2581 GAGAAACATCTTTTCACTGCGGCTGTCTCAGGGGAGAACTCTGATGACGTGGTCCCGTCC 2640
2641 CTGAGCTGCTGCTGACGACACACTCTCTGAGTCCCTCCACTGCTTGTACAGACTCGG 2700
2641 CTGAGCTGCTGCTGACGACACACTCTCTGAGTCCCTCCACTGCTTGTACAGACTCGG 2700
2701 AACAAAACGTTCTTGACAAAGTGAAGGCAATTTGGAAGAAATGGGGATGTGTGTAGAA 2760
2701 AACAAAACGTTCTTGACAAAGTGAAGGCAATTTGGAAGAAATGGGGATGTGTGTAGAA 2760
2761 ACAGACATGGAGCTCTTGTGTGACATTTTCTGCACTTAAATTCAGCCGCCACGTGAAGAAG 2820
2761 ACAGACATGGAGCTCTTGTGTGACATTTTCTGCACTTAAATTCAGCCGCCACGTGAAGAAG 2820
2821 CTTTACAGCTGATTGAGGGCAGGACGACAGATCAACATGGAGCCCAACCATGCTAGTCTG 2880
2821 CTTTACAGCTGATTGAGGGCAGGACGACAGATCAACATGGAGCCCAACCATGCTAGTCTG 2880
2881 TTTAGGTTGGGTCCTGACAGATGCTTATTTGGCAGATTTCTCTCCGTCTCANGGTC 2940
2881 TTTAGGTTGGGTCCTGACAGATGCTTATTTGGCAGATTTCTCTCCGTCTCANGGTC 2940
2941 ACCAGAAACCTGAAAGAGCTGGAACCTTAAGTGGAAATCTCGCTGAGCCACTCTGCACTGAAG 3000
2941 ACCAGAAACCTGAAAGAGCTGGAACCTTAAGTGGAAATCTCGCTGAGCCACTCTGCACTGAAG 3000
3001 AGTCTTTTGAAGCCTCTGAGACGCTCTGCTGCTCTCTGAGAGACCTCGGCTGCTGCTG 3060
3001 AGTCTTTTGAAGCCTCTGAGACGCTCTGCTGCTCTCTGAGAGACCTCGGCTGCTGCTG 3060
3061 TGTGGCCTCAGAGCTGAGGACTGCAAGGACCTTGCCTTTGGGCTGAGAGCCAAACAGACC 3120
3061 TGTGGCCTCAGAGCTGAGGACTGCAAGGACCTTGCCTTTGGGCTGAGAGCCAAACAGACC 3120
3121 CTGACCGAGCTGGACCTGAGCTTCAATGTGTCTCAGGATGCTTGGAGCCAAACACCTTTGC 3180
3121 CTGACCGAGCTGGACCTGAGCTTCAATGTGTCTCAGGATGCTTGGAGCCAAACACCTTTGC 3180
3181 CAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCGACTGCAAGCTGCTGAGTGTGGCCTC 3240
3181 CAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCGACTGCAAGCTGCTGAGTGTGGCCTC 3240
3241 ACCTGTGACTGCTGCCAGGACCTGGCCTCTGCTGCTTGTAGTGCAGCCCGAGCTGAAGGAG 3300
3241 ACCTGTGACTGCTGCCAGGACCTGGCCTCTGCTGCTTGTAGTGCAGCCCGAGCTGAAGGAG 3300
3301 CTAGACCTGACAGCAACAACTTGGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3360
3301 CTAGACCTGACAGCAACAACTTGGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3360
3361 AGGCATCTGCTGCTGCAAACTCATACGCTGGGCTGGACCAAGCACTCTGAGTGTATGAG 3420
3361 AGGCATCTGCTGCTGCAAACTCATACGCTGGGCTGGACCAAGCACTCTGAGTGTATGAG 3420
3421 ATGAGGCAAGGAACTGAGGGCCCTGAGCAGAGAGAACTCTGAGTGTCTCATCTTACAGAGA 3480
3421 ATGAGGCAAGGAACTGAGGGCCCTGAGCAGAGAGAACTCTGAGTGTCTCATCTTACAGAGA 3480
3481 CGGAAACCAAGTGTGATGACCCCTACTGAGGCGCTTGGATACGGGAGAGATGAGTAATAGC 3540
3481 CGGAAACCAAGTGTGATGACCCCTACTGAGGCGCTTGGATACGGGAGAGATGAGTAATAGC 3540
3541 ACATCCTCACTCAAGCGGACAGACTCGGATCAGAGAGGGCGGCTTCCATGTTGCTCAG 3600
3541 ACATCCTCACTCAAGCGGACAGACTCGGATCAGAGAGGGCGGCTTCCATGTTGCTCAG 3600

QY	3601	GCTAATCTCAAACTCTCGACGCTGACCAAGATCTTCCAAATTCGTGAGATTCGAGAGAA	3660
DB	3601	GCTAATCTCAAACTCTCGACGCTGACCAAGATCTTCCAAATTCGTGAGATTCGAGAGAA	3660
QY	3661	AGTCCCGCAGAGGTAGTACCGGTGGAACTCTTGTGGGTGCTCTCTCGCTCTCAAGGG	3720
DB	3661	AGTCCCGCAGAGGTAGTACCGGTGGAACTCTTGTGGGTGCTCTCTCGCTCTCAAGGG	3720
QY	3721	GACCTGCATACGAAGCTTTGGGACCTGACGATGACTCTGGGGCCCCACGGGCTGTG	3780
DB	3721	GACCTGCATACGAAGCTTTGGGACCTGACGATGACTCTGGGGCCCCACGGGCTGTG	3780
QY	3781	GCTACTGAGGTAGTTCACAAAGAAAGAACTTGTACCGAGTTCACTTCCTGTAGCTGGC	3840
DB	3781	GCTACTGAGGTAGTTCACAAAGAAAGAACTTGTACCGAGTTCACTTCCTGTAGCTGGC	3840
QY	3841	TGCTACCGCTGGCCCAACACCGGTCTCTGCTTTGTGATGAGAGAACGCGTGTGAG	3900
DB	3841	TGCTACCGCTGGCCCAACACCGGTCTCTGCTTTGTGATGAGAGAACGCGTGTGAG	3900
QY	3901	ATTGAATCTGTGTGGGACCAAGTTCTCTGGGTGAGATCAACCCACAGACAGCTGGATG	3960
DB	3901	ATTGAATCTGTGTGGGACCAAGTTCTCTGGGTGAGATCAACCCACAGACAGCTGGATG	3960
QY	3961	GTGGCAGGGCTCTGCTGGACATCAAGGCTGAGCCTGSGAGCTGTGGAAGCTGTGACCTC	4020
DB	3961	GTGGCAGGGCTCTGCTGGACATCAAGGCTGAGCCTGSGAGCTGTGGAAGCTGTGACCTC	4020
QY	4021	CCTCACTTTGTGCTCTCAAGGGGGCGATGTGGACACATCCCTGTTCCTCAATATGGCCAC	4080
DB	4021	CCTCACTTTGTGCTCTCAAGGGGGCGATGTGGACACATCCCTGTTCCTCAATATGGCCAC	4080
QY	4081	TTTAAAGAGAGGGATGCTCTCGGAGAGCCAGCCAGCGGTGAGCTGCATCACATGTT	4140
DB	4081	TTTAAAGAGAGGGATGCTCTCGGAGAGCCAGCCAGCGGTGAGCTGCATCACATGTT	4140
QY	4141	CTGGAAGAGAGGGATGCTCTCGGAGAGCCAGCCAGCGGTGAGCTGCATCACATGTT	4200
DB	4141	CTGGAAGAGAGGGATGCTCTCGGAGAGCCAGCCAGCGGTGAGCTGCATCACATGTT	4200
QY	4201	CGCTTCATTCCTGCTGCTGTGGGTGCTCTTCAAGGAGTCTCTGGAATGATGATGATG	4260
DB	4201	CGCTTCATTCCTGCTGCTGTGGGTGCTCTTCAAGGAGTCTCTGGAATGATGATGATG	4260
QY	4261	ACCTTCACCTCTACCTGATCCCAAGTGAAGTCTGCTGCTGGAAGAACTGGAGCTCTGC	4320
DB	4261	ACCTTCACCTCTACCTGATCCCAAGTGAAGTCTGCTGCTGGAAGAACTGGAGCTCTGC	4320
QY	4321	TATCGAAGCCCTGGAAGACCAAGCTGTTCTCGAGTCTTACCTGGCCACTTGGGATCA	4380
DB	4321	TATCGAAGCCCTGGAAGACCAAGCTGTTCTCGAGTCTTACCTGGCCACTTGGGATCA	4380
QY	4381	GGGATCAGGCTCAAGTGAAGAAACAAGAAAGTGAAGTCTGCTGGAGGCTTGGTG	4440
DB	4381	GGGATCAGGCTCAAGTGAAGAAACAAGAAAGTGAAGTCTGCTGGAGGCTTGGTG	4440
QY	4441	AAACCCAGGAGATCTATGCTGCTGCAACTACTCTGATCCCTCCAGCCGCTAGCCGTACCT	4500
DB	4441	AAACCCAGGAGATCTATGCTGCTGCAACTACTCTGATCCCTCCAGCCGCTAGCCGTACCT	4500
QY	4501	TCACCTCTGGATGCCCGCAGTTGCTGCACTTTGTGGACCAAGTATCGAGAGCACTGATA	4560
DB	4501	TCACCTCTGGATGCCCGCAGTTGCTGCACTTTGTGGACCAAGTATCGAGAGCACTGATA	4560
QY	4561	GCCCGAGTGCATCGGTGGAGGTTGTTCTGGACAACTGATGGAAGAGTCTGGAGCCAG	4620
DB	4561	GCCCGAGTGCATCGGTGGAGGTTGTTCTGGACAACTGATGGAAGAGTCTGGAGCCAG	4620
QY	4621	GACAGTACGAGAGGTTGCTGGCTGAGAACACAGGCCCCAGGATGCGGAAGCTGTTC	4680
DB	4621	GACAGTACGAGAGGTTGCTGGCTGAGAACACAGGCCCCAGGATGCGGAAGCTGTTC	4680

QY	4681	AGCTTGAGCCAGTCTCTGGACCGGAAGTGCAAGATGCAATGCACTTACCAAGCCCTGAAGAG	4740
DB	4681	AGCTTGAGCCAGTCTCTGGACCGGAAGTGCAAGATGCAATGCACTTACCAAGCCCTGAAGAG	4740
QY	4741	ACCCATCTCACTCTCATTTATGGAATCTCTGGGAAGGCGCAGCAAAAGAGGACTCTCGCCA	4800
DB	4741	ACCCATCTCACTCTCATTTATGGAATCTCTGGGAAGGCGCAGCAAAAGAGGACTCTCGCCA	4800
QY	4801	CTCAGCAGCTGAAGTATCAACACACAGCCCTTGACCTTTGAGTCTCTGGCTTGACCC	4860
DB	4801	CTCAGCAGCTGAAGTATCAACACACAGCCCTTGACCTTTGAGTCTCTGGCTTGACCC	4860
QY	4861	TTCTTTGGGTCTCAGTCTTCTCTGCAAAAGTGTGCAATCTGCTTGCCTTCAGCA	4920
DB	4861	TTCTTTGGGTCTCAGTCTTCTCTGCAAAAGTGTGCAATCTGCTTGCCTTCAGCA	4920
QY	4921	CTAAAGTAAATGAACCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATG	4980
DB	4921	CTAAAGTAAATGAACCTTTGATGATGATGATGATGATGATGATGATGATGATGATG	4980
QY	4981	CCACAGGGGGCCCCAGTCCAGTGGCCCTAAACAGCACTCTCAGGGAATGTCTCTGGAGCT	5040
DB	4981	CCACAGGGGGCCCCAGTCCAGTGGCCCTAAACAGCACTCTCAGGGAATGTCTCTGGAGCT	5040
QY	5041	GGCAAGACCCCTGACAGCTCTATAGAGCTCTATCTGGTGGCCACAGAGCCCAAGCTAGA	5100
DB	5041	GGCAAGACCCCTGACAGCTCTATAGAGCTCTATCTGGTGGCCACAGAGCCCAAGCTAGA	5100
QY	5101	GGCTTCGGATCCCATCAGGGCGCAAGAGGAATAGGAGGACATGGACCATTTGCCCTC	5160
DB	5101	GGCTTCGGATCCCATCAGGGCGCAAGAGGAATAGGAGGACATGGACCATTTGCCCTC	5160
QY	5161	TGGCTGTGTCAAGAGGTGAGCCCAAAATTTGGGGTTTCAAGCTGGAGGGCCAGCTGGATTC	5220
DB	5161	TGGCTGTGTCAAGAGGTGAGCCCAAAATTTGGGGTTTCAAGCTGGAGGGCCAGCTGGATTC	5220
QY	5221	TTGGCTTTGTACAGGAAGTCTCAAGAGCAAGCAACAGAGTAAAGTGGGAAGGATTT	5280
DB	5221	TTGGCTTTGTACAGGAAGTCTCAAGAGCAAGCAACAGAGTAAAGTGGGAAGGATTT	5280
QY	5281	ATTCAGAAATTAAGAGGATATCACAGCTCTTTAGAAATTTTACTTAAATTTTAAAGAAAT	5340
DB	5281	ATTCAGAAATTAAGAGGATATCACAGCTCTTTAGAAATTTTACTTAAATTTTAAAGAAAT	5340
QY	5341	TTTACAGAAAAACCCCTATAAATTTAAATTTTACTTAAATTTTAAAGAAATTTAAAGAAAT	5400
DB	5341	TTTACAGAAAAACCCCTATAAATTTAAATTTTACTTAAATTTTAAAGAAATTTAAAGAAAT	5400
QY	5401	ACAAAAAGAAAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT	5444
DB	5401	ACAAAAAGAAAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT	5444

RESULT 3
 ABS5497
 ID ABS5497 standard; cdna; 5523 BP.
 XX
 AC ABS5497;
 XX
 XX
 DT 09-JAN-2003 (first entry)
 XX
 DE cDNA encoding human caspase recruitment domain-7 (CARD-7).
 XX
 KW Human; caspase activity; caspase recruitment domain-7; CARD-7;
 KW caspase-1; pseudo-interleukin-1 beta converting enzyme; IL-1beta;
 KW pseudolice; ICEBERG; cell growth; cell death; inflammation;
 KW apoptosis; caspase activation; cancer; follicular lymphoma;
 KW leukoemia; melanoma; colon cancer; lung carcinoma; viral infection;
 KW autoimmune disease; systemic lupus erythematosus; reactive arthritis;
 KW human immunodeficiency virus infection; HIV infection; ALS;
 KW Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
 KW myocardial infarction; stroke; inflammatory disorder; Crohn's disease;
 KW insulin dependent diabetes mellitus; multiple sclerosis; psoriasis;

KW graft rejection; allergic rhinitis; food allergy; conjunctivitis;
 KW glomerular nephritis; cytostatic; virucide; immunosuppressive;
 KW dermatological; nephrotropic; neuroprotective; cardiant; gene; ss.
 XX Homo sapiens.

XX Key Location/Qualifiers
 FH 523..4812
 CDS /*tag= a
 FT /*product= "CARD-7"

XX US2002128219-A1.

XX 12-SEP-2002.

XX 15-AUG-2001; 2001US-0931071.

XX 27-OCT-1999; 99US-0428252.

XX (BERT/) BERTIN J.

XX (ALNE/) ALNEMRI E. S.

XX Bertin J, Alnemri ES;

XX WPI; 2003-028968/02.

XX P-PSDB; ABG71631.

XX Assays for identifying compound that modulates the interaction of
 PT caspase recruitment domain-8 with a CARD-8 ligand or a compound that
 PT modulates activity of CARD-8 -
 XX Disclosure; Fig 1; 49pp; English.

XX The present invention relates to methods of identifying compounds
 CC that regulate caspase activity using caspase recruitment domain-7
 CC (CARD-7) or caspase recruitment domain-8 (CARD-8). In particular, a
 CC method for identifying a compound that modulates the interaction
 CC between CARD-8 and caspase-1, pseudo-interleukin-1 (IL-1) beta
 CC converting enzyme (pseudolICE) or ICEBERG is disclosed. CARD-7 and
 CC CARD-8 molecules are useful as modulating agents in regulating a
 CC variety of cellular processes including cell growth, cell death, and
 CC inflammation. The methods of the invention are useful for identifying
 CC compounds that have the ability to increase/decrease apoptosis, or
 CC comprise the ability to induce caspase activation. The methods are
 CC useful for treating a disorder associated with inappropriate apoptosis
 CC or inappropriate inflammation. The methods are useful for treating
 CC disorders associated with an undesirably low rate of apoptosis such
 CC as cancer (preferably follicular lymphoma, chronic myelogenous
 CC leukaemia, melanoma, colon cancer, lung carcinoma, etc), viral
 CC infections, autoimmune diseases caused by low levels of apoptosis
 CC (e.g. systemic lupus erythematosus, immune-mediated glomerulonephritis,
 CC and arthritis). The methods are also useful for treating disorders with
 CC undesirably high rates of apoptosis such as human immunodeficiency
 CC virus (HIV) infection, Alzheimer's disease, Parkinson's disease,
 CC amyotrophic lateral sclerosis (ALS), retinitis pigmentosa, spinal
 CC muscular atrophy, various forms of cerebellar degeneration, anaemia
 CC associated with chronic disease, aplastic anaemia, chronic neutropenia,
 CC myelodysplastic syndromes, myocardial infarction, stroke, and
 CC various inflammatory disorders (e.g. Crohn's disease, reactive
 CC arthritis, insulin dependent diabetes mellitus, multiple sclerosis,
 CC psoriasis, graft rejection, allergic rhinitis, food allergies,
 CC conjunctivitis, glomerular nephritis, etc). The present sequence
 CC encodes human CARD-7.

XX Sequence 5523 BP; 1350 A; 1529 C; 1481 G; 1163 T; 0 other;

XX Query Match 99.9%; Score 5439.2; DB 25; Length 5523;

XX Best Local Similarity 99.9%; Pred. No. 0;

XX Matches 5441; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GCCCCAGGCGCTGGAGAGTCTGAAGAAACCTGGGAGCCAGAGCCCGGGGCTCACTCT 60

Db 1 GCCCCAGGCGCTGGAGAGTCTGAAGAAACCTGGGAGCCAGAGCCCGGGGCTCACTCT 60

QY 61 GGGTTCTGAAGCCCATTCCTGCTCTCGGGTCTCTCCACCCCACTTCTTCTCAGCCTT 120
 Db |||||
 61 GGGTTCTGAAGCCCATTCCTGCTCTCGGGTCTCTCCACCCCACTTCTTCTCAGCCTT 120

QY 121 GCAGCTCAAGGGTGTATCTCAGGAGTCCAGGACCAGAGAGGGAGAAATCTGAGGAACA 180
 Db |||||
 121 GCAGCTCAAGGGTGTATCTCAGGAGTCCAGGACCAGAGAGGGAGAAATCTGAGGAACA 180

QY 181 CAGAACAGTGAAGGTTGCCACACCCCATCTCCGTCACCAATCTCCCTCAGCCTCAC 240
 Db |||||
 181 CAGAACAGTGAAGGTTGCCACACCCCATCTCCCTCAGCCTCAC 240

QY 241 CCTCCCTGCTGGCCCTGGAGCCCATCCAGGACCTCCCTATCAGCTGACTTCTTCCAGT 300
 Db |||||
 241 CCTCCCTGCTGGCCCTGGAGCCCATCCAGGACCTCCCTATCAGCTGACTTCTTCCAGT 300

QY 301 GTCTTGCAGGCGCTCTGGGCTCTCTCCCTGGCTTTTCTTACCACTCCCTCTCTAT 360
 Db |||||
 301 GTCTTGCAGGCGCTCTGGGCTCTCTCCCTGGCTTTTCTTACCACTCCCTCTCTAT 360

QY 361 CGGCGTCTATCTGTAGTGCCTGGGATTATAAACTGGGTTCCGAATGCTGAATAAGA 420
 Db |||||
 361 CGGCGTCTATCTGTAGTGCCTGGGATTATAAACTGGGTTCCGAATGCTGAATAAGA 420

QY 421 GACGGTAAGAGCCAAAGGCAAGGACAGCACTGTTCTCTGCTGCTGCTGATACCTCAC 480
 Db |||||
 421 GACGGTAAGAGCCAAAGGCAAGGACAGCACTGTTCTCTGCTGCTGATACCTCAC 480

QY 481 CTGGGAAACATCCCCAGACACCTCTTAACTCCGGGACAGAGATGGCTGGGAGACCTGG 540
 Db |||||
 481 CTGGGAAACATCCCCAGACACCTCTTAACTCCGGGACAGAGATGGCTGGGAGACCTGG 540

QY 541 GGCCTGCTGGCTGTTACTTGGAGTTCCTGAAGAGGAGGAGCTGAAGAGTTCAGCTT 600
 Db |||||
 541 GGCCTGCTGGCTGTTACTTGGAGTTCCTGAAGAGGAGGAGCTGAAGAGTTCAGCTT 600

QY 601 CTGCTCGCCAAATAAGCCCACTCCAGGAGCTCTTCGGGTGAGACACCCCGCTCAGC 660
 Db |||||
 601 CTGCTCGCCAAATAAGCCCACTCCAGGAGCTCTTCGGGTGAGACACCTGCTCAGC 660

QY 661 AAGACAGTGGATGAGAGTGGCTCTGTAAGTGGCTGCTGATGAGGAGAGCGGGGCC 720
 Db |||||
 661 AAGACAGTGGATGAGAGTGGCTCTGTAAGTGGCTGCTGATGAGGAGAGCGGGGCC 720

QY 721 TGGGACCTAGCCCTCCATACCTGGGAGAGATGGGCTGAGGCTCACTGTCGCCCAAGCC 780
 Db |||||
 721 TGGGACCTAGCCCTCCATACCTGGGAGAGATGGGCTGAGGCTCACTGTCGCCCAAGCC 780

QY 781 CAGGAAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
 Db |||||
 781 CAGGAAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840

QY 841 TCTCCAGCCAAACCCCACTCCAGCGAGTGTAAATGCTGATGAGGAGAGCGGGCG 900
 Db |||||
 841 TCTCCAGCCAAACCCCACTCCAGCGAGTGTAAATGCTGATGAGGAGAGCGGGCG 900

QY 901 GGGTGCACCCAGGCTCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
 Db |||||
 901 GGGTGCACCCAGGCTCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960

QY 961 CGCTGGAGAGAAATCTCTGCTCTCACTCTCTACCAAGCTCTTCCAGGCTCCCGAGC 1020
 Db |||||
 961 CGCTGGAGAGAAATCTCTGCTCTCACTCTCTACCAAGCTCTTCCAGGCTCCCGAGC 1020

QY 1021 GAGTCTCAAGCCAGGAGTCAACCAACCGCCCCCACTCCACAGCAGTCTCTGGGAGCTGG 1080
 Db |||||
 1021 GAGTCTCAAGCCAGGAGTCAACCAACCGCCCCCACTCCACAGCAGTCTCTGGGAGCTGG 1080

QY 1081 GGATCCCCACCTCAGCCAGGCTTAGCACCCAGAGAGAGGAGGAGGAGGAGGAGGAGG 1140
 Db |||||
 1081 GGATCCCCACCTCAGCCAGGCTTAGCACCCAGAGAGAGGAGGAGGAGGAGGAGGAGG 1140

Qy	1141	CCTCTGATGA	AAACGTCAGG	AAATTTACTACAGAAAT	CAGAGAAAGAGAGAGAGAGAA	1200	
Db	1141	CCTCTGATGA	AAACGTCAGG	AAATTTACTACAGAAAT	CAGAGAAAGAGAGAGAGAGAA	1200	
Qy	1201	TCAGAAAGAG	CGAGGCCCAT	TGGCAGCGGT	GTAGGAACGCCCCACAGGCGCACACC	1260	
Db	1201	TCAGAAAGAG	CGAGGCCCAT	TGGCAGCGGT	GTAGGAACGCCCCACAGGCGCACACC	1260	
Qy	1261	AGCCTCAGC	CCCCACCA	CCCATGGSAGCCTT	CTGTGAGAGAGAGCCCTCTGTTCACA	1320	
Db	1261	AGCCTCAGC	CCCCACCA	CCCATGGSAGCCTT	CTGTGAGAGAGAGCCCTCTGTTCACA	1320	
Qy	1321	TGGCCCTGGA	AAAAATGAG	ATTTTAAACCAAAAT	TTCAACACAGCTGTACTTTCTACAAAGA	1380	
Db	1321	TGGCCCTGGA	AAAAATGAG	ATTTTAAACCAAAAT	TTCAACACAGCTGTACTTTCTACAAAGA	1380	
Qy	1381	CCTCACCC	CAGAGCCAA	GTCCCTGGT	CAAGAGAGCTGGCTGATATGTGGAGGAG	1440	
Db	1381	CCTCACCC	CAGAGCCAA	GTCCCTGGT	CAAGAGAGCTGGCTGATATGTGGAGGAG	1440	
Qy	1441	AATCGAGCA	CATTTAA	TTGAGATCAG	AGACTTATTTGGCCCCAGGCTTGGATACCCAGAA	1500	
Db	1441	AATCGAGCA	CATTTAA	TTGAGATCAG	AGACTTATTTGGCCCCAGGCTTGGATACCCAGAA	1500	
Qy	1501	CCTCGCAT	GTATCTG	CAGGGGCTG	TGGAAATTTGGGAAGTCAACACTGGCCAGGCAG	1560	
Db	1501	CCTCGCAT	GTATCTG	CAGGGGCTG	TGGAAATTTGGGAAGTCAACACTGGCCAGGCAG	1560	
Qy	1561	GTGAAGAA	CGCTGGG	GAGGCCAG	CTGTATGGGACCGCTTCCAGCATGTCTTCTAC	1620	
Db	1561	GTGAAGAA	CGCTGGG	GAGGCCAG	CTGTATGGGACCGCTTCCAGCATGTCTTCTAC	1620	
Qy	1621	TTCAGCTG	CAGAGCTG	GGCCCACT	CGCAAGTGGTGA	GTCTCGCTGAGCTCATTCGGAAAA	1680
Db	1621	TTCAGCTG	CAGAGCTG	GGCCCACT	CGCAAGTGGTGA	GTCTCGCTGAGCTCATTCGGAAAA	1680
Qy	1681	GATGGGA	CAGCACT	CCGGCTCC	ATATAGACAGATCTGTCTAGGCAGAGCGCTGCTC	1740	
Db	1681	GATGGGA	CAGCACT	CCGGCTCC	ATATAGACAGATCTGTCTAGGCAGAGCGCTGCTC	1740	
Qy	1741	TTCAATCT	CGATGCT	GATGAC	CGAGATGGGTCTTGCAGGAGCCGAGTCTTGAGCTC	1800	
Db	1741	TTCAATCT	CGATGCT	GATGAC	CGAGATGGGTCTTGCAGGAGCCGAGTCTTGAGCTC	1800	
Qy	1801	TGCTGCA	CTGGAC	CCAGCCAG	CCGCGGATGCACTGTGGGCAGTGTCTGGGGAAA	1860	
Db	1801	TGCTGCA	CTGGAC	CCAGCCAG	CCGCGGATGCACTGTGGGCAGTGTCTGGGGAAA	1860	
Qy	1861	ACTATACT	CCCCAGG	CATCCTT	CGATCAGCGCTCGGACCA	CAGCTCTGCAGAACCTC	1920
Db	1861	ACTATACT	CCCCAGG	CATCCTT	CGATCAGCGCTCGGACCA	CAGCTCTGCAGAACCTC	1920
Qy	1921	ATTCCCTT	CTTGGAG	CGAGCCAG	TGGTAGAGTCTCTGGGTTCTCTGAGTCCAGCAGG	1980	
Db	1921	ATTCCCTT	CTTGGAG	CGAGCCAG	TGGTAGAGTCTCTGGGTTCTCTGAGTCCAGCAGG	1980	
Qy	1981	AAGGAAT	ATTTCTA	CAGATATTT	CAACAGATGAAAGCAAGCAAT	TAGAGCCTTTAGTTG	2040
Db	1981	AAGGAAT	ATTTCTA	CAGATATTT	CAACAGATGAAAGCAAGCAAT	TAGAGCCTTTAGTTG	2040
Qy	2041	GTCAAAAT	CAAAAG	AGCTCTG	GGCCCTGTGTCTTGTG	CCCTGGGTGTCTGGCTGGCC	2100
Db	2041	GTCAAAAT	CAAAAG	AGCTCTG	GGCCCTGTGTCTTGTG	CCCTGGGTGTCTGGCTGGCC	2100
Qy	2101	TGCACTTC	CTTGAG	CGAGCCAG	TGAGTCTCTGGGTTCTCTGAGTCCAGCAGG	1980	
Db	2101	TGCACTTC	CTTGAG	CGAGCCAG	TGAGTCTCTGGGTTCTCTGAGTCCAGCAGG	1980	
Qy	2161	ACCACAA	CCCTCTGT	CTACATTA	CTTGC	CCAGGCTCTTCAAGCTCAGCCATTTGGGACCC	2220
Db	2161	ACCACAA	CCCTCTGT	CTACATTA	CTTGC	CCAGGCTCTTCAAGCTCAGCCATTTGGGACCC	2220
Qy	2221	CAGCTCAG	AGACCTCT	GTCTCTG	GGCTGTGAGG	GCATCTGGCAAAAAAGACCCCTTTTC	2280

2221	CAGCTCAGAGACCTTCGCTCTCTGCGCTGTGAGGGCATCTGCCAAAAAAGACCCCTTTTC	2280
2281	AGTCAGAGATGACCTCAGGAAGCATGGGTAGATGGGGCCATCATCTCCACCTTCTTGAAG	2340
2281	AGTCAGAGATGACCTCAGGAAGCATGGGTAGATGGGGCCATCATCTCCACCTTCTTGAAG	2340
2341	ATGGGTATTCTTCAAGAGACCCCATCCCTCTGAGCTACAGCTTCACTTCACTCTGTGTTTC	2400
2401	CHAGAGTTCCTTTGAGCAATGTCTTATGCTGAGGATGAGAAGGGGAGAGGTTAAACAT	2460
2401	CHAGAGTTCCTTTGAGCAATGTCTTATGCTGAGGATGAGAAGGGGAGAGGTTAAACAT	2460
2461	TCTAATTGTCATCATAGATTTTGAAAGAAGCGCTAGAACATATGGAAATACATGGCCTGTTT	2520
2461	TCTAATTGTCATCATAGATTTTGAAAGAAGCGCTAGAACATATGGAAATACATGGCCTGTTT	2520
2521	GGGGCATCAACACACGTTTTCTCAAGGAGAGAACTGATGAGTGTAGAGGGGAAGAGAGATG	2580
2521	GGGGCATCAACACACGTTTTCTCAAGGAGAGAACTGATGAGTGTAGAGGGGAAGAGAGATG	2580
2581	GAGAACATCTTTTCACTGCGCGCTGCTCAGGGGAGAACTGATGAGTGTAGAGGGTCCCGTCC	2640
2581	GAGAACATCTTTTCACTGCGCGCTGCTCAGGGGAGAACTGATGAGTGTAGAGGGTCCCGTCC	2640
2641	CTGACGCTGCTGCTGACGCCACACTCTCTGAGGTCCCTCCACTGCTGTAGAGACTCGG	2700
2641	CTGACGCTGCTGCTGACGCCACACTCTCTGAGGTCCCTCCACTGCTGTAGAGACTCGG	2700
2701	AACAAAAAGTTCCTTGACACAAAGTATGGCCCATTTTCGAGAAATGGGCATGTGTGTAGAA	2760
2701	AACAAAAAGTTCCTTGACACAAAGTATGGCCCATTTTCGAGAAATGGGCATGTGTGTAGAA	2760
2761	ACAGACATGGAGCTCTTAGTGTGCACTTCTTGCACTTAAATTTAGCGCGCCAAGTGAAGAG	2820
2761	ACAGACATGGAGCTCTTAGTGTGCACTTCTTGCACTTAAATTTAGCGCGCCAAGTGAAGAG	2820
2821	CTTCAGCTGATTGAGGGCAGCAGACAGATCAACATGAGGCCCAACCATGGTAGTCTCGT	2880
2821	CTTCAGCTGATTGAGGGCAGCAGACAGATCAACATGAGGCCCAACCATGGTAGTCTCGT	2880
2881	TTCAGTGGGTCCCAGTCACAGATGCCCTATTGGCAGATTTCTTTCTCCGTCTCAAGGTC	2940
2881	TTCAGTGGGTCCCAGTCACAGATGCCCTATTGGCAGATTTCTTTCTCCGTCTCAAGGTC	2940
2941	ACAGAAACCTGAGGAGCTGGACTAAGTGGAAAACTCGCTGAGCCACTCTGCGAGTGAAG	3000
2941	ACAGAAACCTGAGGAGCTGGACTAAGTGGAAAACTCGCTGAGCCACTCTGCGAGTGAAG	3000
3001	AGTCTTTGTAAGACCCCTCAGAGGCCCTCGTGCCTCTCTGGAGACCTCGCGTTGCTGGCC	3060
3001	AGTCTTTGTAAGACCCCTCAGAGGCCCTCGTGCCTCTCTGGAGACCTCGCGTTGCTGGCC	3060
3061	TGTGGCCTCACAGCTGAGGACTGCAAGGACCTTGCCCTTTGGGCTGAGAGCCAAACAGACC	3120
3061	TGTGGCCTCACAGCTGAGGACTGCAAGGACCTTGCCCTTTGGGCTGAGAGCCAAACAGACC	3120
3121	CTGACCGAGCTGGACTGAGCTTCAATGTGTCTACGGATGCTGGAGCCMAACACCTTTGC	3180
3121	CTGACCGAGCTGGACTGAGCTTCAATGTGTCTACGGATGCTGGAGCCMAACACCTTTGC	3180
3181	CAGAGACTCAGACAGCCAGCTGCAAGCTACGCACTGCTGCTGAGTGTGAGTGTGGCCCTC	3240
3181	CAGAGACTCAGACAGCCAGCTGCAAGCTACGCACTGCTGCTGAGTGTGAGTGTGGCCCTC	3240
3241	ACGCTCTGACTGCTGCCAGGACCTGGCCTCTGTGCTTAGTGCAGCCCGCCAGCTGAAGAG	3300
3241	ACGCTCTGACTGCTGCCAGGACCTGGCCTCTGTGCTTAGTGTGAGTGTGAGTGTGGCCCTC	3300
3301	CTAGACCTCGACAGAAACCTGGATGACGTTGGCGTGGCACTGCTGTGAGGGGCTC	3360

Db 3301 CTAGACCTGCAGCAGAACAACTCGATGACGTTGGCGTGCAGCTGCTGTGAGGGGCTC 3360
 Qy 3361 AGGCATCCCTGCTGCAAACTCATACGCTGGGCTGGAACAGACAACTCTGAGTGATGAG 3420
 Db 3361 AGGCATCCCTGCTGCAAACTCATACGCTGGGCTGGAACAGACAACTCTGAGTGATGAG 3420
 Qy 3421 ATGAGGCAGGAACTGAGGGCCCTGGAGCAGGAGAACTCTCAGCTGCTCATCTTCAGCAGA 3480
 Db 3421 ATGAGGCAGGAACTGAGGGCCCTGGAGCAGGAGAACTCTCAGCTGCTCATCTTCAGCAGA 3480
 Qy 3481 CGGAAACCAAGTGTGATGACCCCTTA CTGAGGGCCTGGATACGGGAGAGATGATGATAGC 3540
 Db 3481 CGGAAACCAAGTGTGATGACCCCTTA CTGAGGGCCTGGATACGGGAGAGATGATGATAGC 3540
 Qy 3541 ACATCCTCACTCAAGCGGAGAGATCGATCAGAGAGGGCGCTTCCCATGTTGCTCAG 3600
 Db 3541 ACATCCTCACTCAAGCGGAGAGATCGATCAGAGAGGGCGCTTCCCATGTTGCTCAG 3600
 Qy 3601 GCTAATCTCAAACTCCTGGACGTGACAAAGATCTTCCCAATTGCTGAGATTGCAGAGGAA 3660
 Db 3601 GCTAATCTCAAACTCCTGGACGTGACAAAGATCTTCCCAATTGCTGAGATTGCAGAGGAA 3660
 Qy 3661 AGCTCCCAAGAGTAGTA CCGGTGGAAC TCTTGTGCGTGCCTTCTCCTGCCCTCAAGGG 3720
 Db 3661 AGCTCCCAAGAGTAGTA CCGGTGGAAC TCTTGTGCGTGCCTTCTCCTGCCCTCAAGGG 3720
 Qy 3721 GACCTGCATACGAAGCTTTGGGACTGACGATGACTCTGGGGCCCAAGGGGCTGTG 3780
 Db 3721 GACCTGCATACGAAGCTTTGGGACTGACGATGACTCTGGGGCCCAAGGGGCTGTG 3780
 Qy 3781 GCTACTGAGTAGTTGACAAAGAAAGAACTTGTACCGAGTTCACTTCCCTGTAGCTGGC 3840
 Db 3781 GCTACTGAGTAGTTGACAAAGAAAGAACTTGTACCGAGTTCACTTCCCTGTAGCTGGC 3840
 Qy 3841 TCTACCGCTGGCCCAACACGGGTCTCTGCTTTGTGATGAGAGACGGTGACCGTTGAG 3900
 Db 3841 TCTACCGCTGGCCCAACACGGGTCTCTGCTTTGTGATGAGAGACGGGTGACCGTTGAG 3900
 Qy 3901 ATTGAATCTGTGTGGGACCAAGTTCCTGGGTGAGATCAACCCACAGACAGCTGGATG 3960
 Db 3901 ATTGAATCTGTGTGGGACCAAGTTCCTGGGTGAGATCAACCCACAGACAGCTGGATG 3960
 Qy 3961 GTGGCAGGGCCTCTGTGTCACATCAAGCTGAGCCTGAGCTGTGGAAGCTGTGCACCTC 4020
 Db 3961 GTGGCAGGGCCTCTGTGTCACATCAAGCTGAGCCTGAGCTGTGGAAGCTGTGCACCTC 4020
 Qy 4021 CCTCACTTTGTGCTCTCAAGGGGCCATGTGGAACATCCCTGTTCCAAATGGCCAC 4080
 Db 4021 CCTCACTTTGTGCTCTCAAGGGGCCATGTGGAACATCCCTGTTCCAAATGGCCAC 4080
 Qy 4081 TTTAAAGAGGGGATGCTCTTGGAGAGCCAGCCAGGGTGGAGCTGCATCACATAGTT 4140
 Db 4081 TTTAAAGAGGGGATGCTCTTGGAGAGCCAGCCAGGGTGGAGCTGCATCACATAGTT 4140
 Qy 4141 CTGGAAAAACCCAGCTTCTCCCTTTGGAGTCTCTCTGAAAATGATCAATAATGCCCTG 4200
 Db 4141 CTGGAAAAACCCAGCTTCTCCCTTTGGAGTCTCTCTGAAAATGATCAATAATGCCCTG 4200
 Qy 4201 CGCTTCATTTCCGTCA CCTCTGTGTGTTGCTTTTACCA CCGGTCCATCCTCAGGAAGTC 4260
 Db 4201 CGCTTCATTTCCGTCA CCTCTGTGTGTTGCTTTTACCA CCGGTCCATCCTCAGGAAGTC 4260
 Qy 4261 ACCTTCCA CTTACTGATCCCAAGTACTGCTCCATTTCCGGAAGAACTGGAGCTCTGC 4320
 Db 4261 ACCTTCCA CTTACTGATCCCAAGTACTGCTCCATTTCCGGAAGAACTGGAGCTCTGC 4320
 Qy 4321 TATCGAAGCCCTGGAGAAACCAAGTCTCGAGTTCTAGCTTGGCCCACTTGGGATCA 4380
 Db 4321 TATCGAAGCCCTGGAGAAACCAAGTCTCGAGTTCTAGCTTGGCCCACTTGGGATCA 4380
 Qy 4381 GGGATCAGGCTGCAAGTGAAGAAACAAAGATGAGACTCTGTGTGGAGGCGCTTGGTG 4440
 Db 4381 GGGATCAGGCTGCAAGTGAAGAAACAAAGATGAGACTCTGTGTGGAGGCGCTTGGTG 4440

Qy 4441 AAACGAGAGATCTCATGCTGCAACTACTCTGATCCCTCAGCCCGCATAGCCGTACT 4500
 Db 4441 AAACGAGAGATCTCATGCTGCAACTACTCTGATCCCTCAGCCCGCATAGCCGTACT 4500
 Qy 4501 TCACCTCTGGATGCCCCCGAGTTGCTGCACTTTGTGGAACAGTATCGAGAGAGCTGATA 4560
 Db 4501 TCACCTCTGGATGCCCCCGAGTTGCTGCACTTTGTGGAACAGTATCGAGAGAGCTGATA 4560
 Qy 4561 GCCCGAGTGCATCGGTGGAGGTTGCTTGGACAAAACCTGCAATGGACAGCTGAGCCAG 4620
 Db 4561 GCCCGAGTGCATCGGTGGAGGTTGCTTGGACAAAACCTGCAATGGACAGCTGAGCCAG 4620
 Qy 4621 GAGCAGTACGAGAGGCTGCTGCTGAGAAACAGAGGCCAGCAGATCGGAAGCTGTTTC 4680
 Db 4621 GAGCAGTACGAGAGGCTGCTGCTGAGAAACAGAGGCCAGCAGATCGGAAGCTGTTTC 4680
 Qy 4681 AGCTTGGAGCCAGTCTGGGACCGGAAGTCAAGATGAGACTCTACCAAGCCCTGAAGGAG 4740
 Db 4681 AGCTTGGAGCCAGTCTGGGACCGGAAGTCAAGATGAGACTCTACCAAGCCCTGAAGGAG 4740
 Qy 4741 ACCCATCTCACTCATTTATGGAACCTCTGGAGAGGCGACAAAAGGGACTCTCTGCCA 4800
 Db 4741 ACCCATCTCACTCATTTATGGAACCTCTGGAGAGGCGACAAAAGGGACTCTCTGCCA 4800
 Qy 4801 CTCAGCAGCTCAAGTATCAACACAGCCCTTGACCCCTTGAGTCTCTGGCTTGACCC 4860
 Db 4801 CTCAGCAGCTCAAGTATCAACACAGCCCTTGACCCCTTGAGTCTCTGGCTTGACCC 4860
 Qy 4861 TTTCTTGGGTCTCAGTTCTTCTCTGCAAAACAGATTGCCATCTGTTTGGCTTCCAGCA 4920
 Db 4861 TTTCTTGGGTCTCAGTTCTTCTCTGCAAAACAGATTGCCATCTGTTTGGCTTCCAGCA 4920
 Qy 4921 CTAAGTAAATGGAATTTGATGATGCTTCTGCTGGCATTTGTTCCATGCCAGGATG 4980
 Db 4921 CTAAGTAAATGGAATTTGATGATGCTTCTGCTGGCATTTGTTCCATGCCAGGATG 4980
 Qy 4981 CCACAGGGGGCCCAAGTCCAGTGGCTTAAACAGATCTCAGGGAATGTCATCTGGAGCT 5040
 Db 4981 CCACAGGGGGCCCAAGTCCAGTGGCTTAAACAGATCTCAGGGAATGTCATCTGGAGCT 5040
 Qy 5041 GGCAAGACCCCTGCAGACCTCATAGAGCTCATCTGTGGCCACAGCAGCAAGCCTAGA 5100
 Db 5041 GGCAAGACCCCTGCAGACCTCATAGAGCTCATCTGTGGCCACAGCAGCAAGCCTAGA 5100
 Qy 5101 GCCCTCCGGATCCCATCCAGGGCAAGAGAAATAGGAGGCACTGGAAACCATTTGCCCTC 5160
 Db 5101 GCCCTCCGGATCCCATCCAGGGCAAGAGAAATAGGAGGCACTGGAAACCATTTGCCCTC 5160
 Qy 5161 TGGCTGTCTCAGGGTCCAGCCCAAAATTTGGGTTTACGCTGGGAGGCCAGCTGGATTC 5220
 Db 5161 TGGCTGTCTCAGGGTCCAGCCCAAAATTTGGGTTTACGCTGGGAGGCCAGCTGGATTC 5220
 Qy 5221 TTGGCTTTTGTACAGGAAGTCTCAAGAGCAAGCAACAGAGTAAAGTGAAGGAAGTTT 5280
 Db 5221 TTGGCTTTTGTACAGGAAGTCTCAAGAGCAAGCAACAGAGTAAAGTGAAGGAAGTTT 5280
 Qy 5281 ATTCAAAAAATAAGGAGTATCAAGCTCTTTTGAATTTTCTAGCAGGCTTTCCAGTT 5340
 Db 5281 ATTCAAAAAATAAGGAGTATCAAGCTCTTTTGAATTTTCTAGCAGGCTTTCCAGTT 5340
 Qy 5341 TTTTACAGAAAAACCCCTATAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 5400
 Db 5341 TTTTACAGAAAAACCCCTATAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 5400
 Qy 5401 ACAAAAAAGAAAAATGAAAAATAAGGAATGAAGTTTACCTAC 5444
 Db 5401 ACAAAAAAGAAAAATGAAAAATAAGGAATGAAGTTTACCTAC 5444

RESULT 4
 AAL47127
 ID AAL47127 standard; DNA; 5100 BP.

XX	AA47127;	Db	393	GCTGAGGTCACTGTGGCGCCCAAGCCAGGAAGGGGAGGCGACCTCTCCTCATTTCCCTA	452
AC		Qy	816	CAGCCCAAGTGAACCCCACTGGGGTCTCCAGCGCAACCCACCTCCACCGAGTGTAA	875
XX		Db	453	CAGCCCAAGTGAACCCCACTGGGGTCTCCAGCGCAACCCACCTCCACCGAGTGTAA	512
DT		Qy	876	GCCCTGGATCCATGAATTGCGGGGGGGTGACCCAGGGCTCAGAGAGAAGGTTTGG	935
XX		Db	513	GCCCTGGATCCATGAATTGCGGGGGGGTGACCCAGGGCTCAGAGAGAAGGTTTGG	572
DE	Pyrin domain containing protein NALP1-hs coding sequence.	Qy	936	ACAGTGGCTCAGACATCTGGACCGCTGGAGAGAAATCTCTGCTCACTCTCTACCA	995
XX		Db	573	ACAGTGGCTCAGACATCTGGACCGCTGGAGAGAAATCTCTGCTCACTCTCTACCA	632
XX		Qy	996	AGCTCTTCAAGCTCCCCAGACCATGAGTCTCCAGCCAGGAGTCAACCAACGCCCCAC	1055
XX		Db	633	AGCTCTTCAAGCTCCCCAGACCATGAGTCTCCAGCCAGGAGTCAACCAACGCCCCAC	692
PN	WO200240668-A2.	Qy	1056	ATCCACAGAGTGTGGGGAGCTGGGGATCCCACTCAGCCAGCTAGACCCAGAGA	1115
XX		Db	693	ATCCACAGAGTGTGGGGAGCTGGGGATCCCACTCAGCCAGCTAGACCCAGAGA	752
XX		Qy	1116	GCAGGAGCTCTGGGACCCCAATGSCCTCTGATGAAACGTCAGGAATTTACTACACAGA	1175
XX		Db	753	GCAGGAGCTCTGGGACCCCAATGSCCTCTGATGAAACGTCAGGAATTTACTACACAGA	812
PA	(APOT-) APOTECH RES & DEV LTD.	Qy	1176	AATCAGAGAAAGAGAGAGAGAAATCAGAGAAAGCGAGGCCCGCCATGGGAGCGGTGT	1235
XX		Db	813	AATCAGAGAAAGAGAGAGAGAAATCAGAGAAAGCGAGGCCCGCCATGGGAGCGGTGT	872
PI	Tschopp J, Martinon F;	Qy	1236	AGGAA CGGCCCCCAGGCGCACACAGGCTTACAGCCCAACCAACCCATGGAGCCCTTC	1295
XX		Db	873	AGGAA CGGCCCCCAGGCGCACACAGGCTTACAGCCCAACCAACCCATGGAGCCCTTC	932
DR	WPI; 2002-427093/45.	Qy	1296	TGTGAGAGAGCCCTCTGTTCCACATGSCCTGGAGAAATGAGGATTTTAAACCAAAAT	1355
DR	P-PSDB; AA017855.	Db	933	TGTGAGAGAGCCCTCTGTTCCACATGSCCTGGAGAAATGAGGATTTTAAACCAAAAT	992
XX		Qy	1356	CACACAGCTGTACTTCTACAAAGACCTCACCCAGAGCCAAAGATCCCTGGTCAAGAG	1415
XX		Db	993	CACACAGCTGTACTTCTACAAAGACCTCACCCAGAGCCAAAGATCCCTGGTCAAGAG	1052
CC	The present invention relates the DNA and their encoded proteins, where	Qy	1416	AAGCTGGCTGATTTATGGAGGAGAAATCGAGGACATTTAAATGAGATCAGAGACTTAT	1475
CC	the proteins contain at least one PYD (pyrin) domain. These can be used	Db	1053	AAGCTGGCTGATTTATGGAGGAGAAATCGAGGACATTTAAATGAGATCAGAGACTTAT	1112
CC	to treat diseases associated with impaired intracellular signal	Qy	1476	TGGCCAGGCTGGATATCCCAAGACCTCGCATAGTCTACTGAGGGGGCTGTGGAT	1535
CC	transduction, particularly inflammation such as psoriasis,	Db	1113	TGGCCAGGCTGGATATCCCAAGACCTCGCATAGTCTACTGAGGGGGCTGTGGAT	1172
CC	arteriosclerosis, bacterial or viral infections (particularly meningitis	Qy	1536	TGGGAAGTCAACACTGGCGAGGCTGAAGGAGCCCTGGGGAGAGGCGACCTGTATGG	1595
CC	and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,	Db	1173	TGGGAAGTCAACACTGGCGAGGCTGAAGGAGCCCTGGGGAGAGGCGACCTGTATGG	1232
CC	sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's	Qy	1596	GGACCGCTTCCAGCATGTCTTCTACTTCAGCTGACAGAGCTGGCCCACTCCAGGTTGT	1655
CC	the invention.	Db	1233	GGACCGCTTCCAGCATGTCTTCTACTTCAGCTGACAGAGCTGGCCCACTCCAGGTTGT	1292
CC		Qy	1656	GAGTCTGCTGAGCTCATCGGAAAGATGGGACACGCACTCCGGCTCCCATATAGACAT	1715
CC		Db	1293	GAGTCTGCTGAGCTCATCGGAAAGATGGGACACGCACTCCGGCTCCCATATAGACAT	1352
CC		Qy	1716	CCTGTCTAGGCGAGAGCTGTCTTCTATCTCCATGCTGTAGATGAGCCAGGATGGGT	1775
CC		Db	1353	CCTGTCTAGGCGAGAGCTGTCTTCTATCTCCATGCTGTAGATGAGCCAGGATGGGT	1412
CC		Qy	1776	CTTTCAGGAGCGAGTCTGTAGCTCTGTCTGCACTGGAGCCACGACAGCCGCGATGC	1835
CC		Db	1413	CTTTCAGGAGCGAGTCTGTAGCTCTGTCTGCACTGGAGCCACGACAGCCGCGATGC	1472
CC		Qy	1836	ACTGTGGGCACTTTGCTGGGAAAACTATATCTTCCGAGGATCTCTTCTGATCAACGGC	1895
CC		Db	1473	ACTGTGGGCACTTTGCTGGGAAAACTATATCTTCCGAGGATCTCTTCTGATCAACGGC	1532

XX	Query Match	89.6%; Score 4879.2; DB 24; Length 5100;
AC	Best Local Similarity	99.4%; Pred. No. 0;
DT	Matches 4896; Conservative	0; Mismatches 28; Indels 0; Gaps 0;
XX		
DE	516	GACAGAGATGGCTGGCGGAGCTGGGCGCCCTGGGCTGTACTTGGAGTTCTCTGAAGAA
XX	513	GAAACAAGTGGCTGGCGGAGCCCTGGGCGCCCTGGGCTGTACTTGGAGTTCTCTGAAGAA
XX	576	GGAGGAGCTGAAGGAGTTCCAGCTTCTGCTCGCCCAATAAGCGCACTCCAGGAGCTCTTC
XX	213	GGAGGAGCTGAAGGAGTTCCAGCTTCTGCTCGCCCAATAAGCGCACTCCAGGAGCTCTTC
XX	636	GGGTGAGACACCCGCTCAGCCAGAGAGACAGTGGGATGGAGTGGGCTCGTACCTGGT
XX	273	GGGTGAGACACCCGCTCAGCCAGAGAGACAGTGGGATGGAGTGGGCTCGTACCTGGT
XX	696	GGCTCAGTATGGGAGCAGCGGCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGG
XX	333	GGCTCAGTATGGGAGCAGCGGCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGG
XX	756	GCTGAGTCACTGTGGGCCCAAGCCAGGAAGGGGAGGCGCACTCTCCCTCATTTCCCTA

1896 TCGGACCAAGCTCTGACAGAACTCTATCTCTTTTGGAGCAGGACGCTTGGGTAGAGGT 1955
 1533 TCGGACCAAGCTCTGACAGAACTCTATCTCTTTTGGAGCAGGACGCTTGGGTAGAGGT 1592
 1956 CTTGGGGTCTCTGAGTCCAGCAGGAAGAAATATTTCTACAGATATTTTACAGATGAAG 2015
 1593 CTTGGGGTCTCTGAGTCCAGCAGGAAGAAATATTTCTACAGATATTTTACAGATGAAG 1652
 2016 GCAAGCAATTAGAGCCTTTAGGTGGTCAAAATCAAAAGAGCTCTGGGCCCTGTGTCT 2075
 1653 GCAAGCAATTAGAGCCTTTAGGTGGTCAAAATCAAAAGAGCTCTGGGCCCTGTGTCT 1712
 2076 TGTGCCCTGGGTCTCTGCTGGCTGGCTCTGCTGATGACAGCAGATGAAGCGGAAGGA 2135
 1713 TGTGCCCTGGGTCTCTGCTGGCTGGCTCTGCTGATGACAGCAGATGAAGCGGAAGGA 1772
 2136 AAAATCAGCTGACCTTCAAGACCAACCAACCTCTCTCTACATTTACCTTTGCCAGGC 2195
 1773 AAAATCAGCTGACCTTCAAGACCAACCAACCTCTCTCTACATTTACCTTTGCCAGGC 1832
 2196 TCTCCAGCTCAGCCATTTGGGACCCAGCTCAGAGACCTCTGCTCTCTGGCTGTGAGGG 2255
 1833 TCTCCAGCTCAGCCATTTGGGACCCAGCTCAGAGACCTCTGCTCTCTGGCTGTGAGGG 1892
 2256 CATCTGGCAAAAAAGACCCCTTTTCAATCCAGATGACCTCAGAGAGCTGGGTAGATGG 2315
 1893 CATCTGGCAAAAAAGACCCCTTTTCAATCCAGATGACCTCAGAGAGCTGGGTAGATGG 1952
 2316 GGCATCATCTCCACCTCTTGAAGATGGGTATTTTCAAGAGCACCCCATCTCTGAG 2375
 1953 GGCATCATCTCCACCTCTTGAAGATGGGTATTTTCAAGAGCACCCCATCTCTGAG 2012
 2376 CTACAGCTTCAATCAGCTCTGTTTCCAAGATCTTTTGGAGCAATGTCTATGTCTTGA 2435
 2013 CTACAGCTTCAATCAGCTCTGTTTCCAAGATCTTTTGGAGCAATGTCTATGTCTTGA 2072
 2436 GGATGAGAGGGGAGAGGTAAACATTTCAATGTCATAGATTTTGAAGACGCTAGA 2495
 2073 GGATGAGAGGGGAGAGGTAAACATTTCAATGTCATAGATTTTGAAGACGCTAGA 2132
 2496 AGCATATGAAATACATGAGCTGTTTGGGACATCAACACACGTTCTATTTGGGCTGTT 2555
 2133 AGCATATGAAATACATGAGCTGTTTGGGACATCAACACACGTTCTATTTGGGCTGTT 2192
 2556 AAGTGAATGAGGGGAGAGAGATGGAGAAACATCTTTTCACTGCGGCTGTCTCAGGGAG 2615
 2193 AAGTGAATGAGGGGAGAGAGATGGAGAAACATCTTTTCACTGCGGCTGTCTCAGGGAG 2252
 2616 GAACCTGATGACAGTGGGTCCCGTCCCTGAGCTGCTGTCGACGACACCTCTCTGGAGTC 2675
 2253 GAACCTGATGACAGTGGGTCCCGTCCCTGAGCTGCTGTCGACGACACCTCTCTGGAGTC 2312
 2676 CTTCCACATGCTTTGACGAGCTCGGAACAAAACGTTCTTGGACACAGTGAATGGCCATTT 2735
 2313 CTTCCACATGCTTTGACGAGCTCGGAACAAAACGTTCTTGGACACAGTGAATGGCCATTT 2372
 2736 CGAAGAAATGGGCAATGTGTGAGAAACACAGATGGAGCTCTTAGTGTGACATTTCTGCAT 2795
 2373 CGAAGAAATGGGCAATGTGTGAGAAACACAGATGGAGCTCTTAGTGTGACATTTCTGCAT 2432
 2796 TAAATTCAGCCGCCAGTGAAGAGCTTCAGCTGATTGAGGGCAGGACACAGATCAAC 2855
 2433 TAAATTCAGCCGCCAGTGAAGAGCTTCAGCTGATTGAGGGCAGGACACAGATCAAC 2492
 2856 ATGGAGCCCAACCATGGTATGCTGTTTCAAGTGGGTCCAGTCAAGATGCTTATTTGGCA 2915
 2493 ATGGAGCCCAACCATGGTATGCTGTTTCAAGTGGGTCCAGTCAAGATGCTTATTTGGCA 2552
 2916 GATTCTCTTCTCGTCTCAAGGTCAACAGAAACCTTGAAGAGCTGGACCTAAGTGGAA 2975
 2553 GATTCTCTTCTCGTCTCAAGGTCAACAGAAACCTTGAAGAGCTGGACCTAAGTGGAA 2612

2976 CTGCTGAGCCACTCTGACAGTGAAGAGTCTTTTGAAGACCTTGAGACCCCTCTGCTGCTCT 3035
 2613 CTGCTGAGCCACTCTGACAGTGAAGAGTCTTTTGAAGACCTTGAGACCCCTCTGCTGCTCT 2672
 3036 CTTGGAGACCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 3095
 2673 CTTGGAGACCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 2732
 3096 CTTTGGGCTGAGAGCAACACAGACCTTGCAGGAGCTGGACCTGAGCTTCAATGTGCTCAC 3155
 2733 CTTTGGGCTGAGAGCAACACAGACCTTGCAGGAGCTGGACCTGAGCTTCAATGTGCTCAT 2792
 3156 GGATCTGGAGCCAAACACCTTTGCCAGAGACTGAGACAGCCGAGCTCAAGCTACAGCG 3215
 2793 GGATCTGGAGCCAAACACCTTTGCCAGAGACTGAGACAGCCGAGCTCAAGCTACAGCG 2852
 3216 ACTGAGCTGGTCAAGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 3275
 2853 ACTGAGCTGGTCAAGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 2912
 3276 TAGTSCCAGCCCGACCTGAAAGGAGCTAGACCTGCAGCAGAAACCTTGGATGAGCTTGG 3335
 2913 TAGTSCCAGCCCGACCTGAAAGGAGCTAGACCTGCAGCAGAAACCTTGGATGAGCTTGG 2972
 3336 CGTGCAGCTGCTCTGTGAGGGGCTCAGGCATCTGCTTGCCTGCAAACTCATACGCTTGGGCT 3395
 2973 CGTGCAGCTGCTCTGTGAGGGGCTCAGGCATCTGCTTGCCTGCAAACTCATACGCTTGGGCT 3032
 3396 GGACACAGAACTCTGAGTGTAGATGAGCAGGAACTGAGGGCCCTTGGAGCAGGAGAA 3455
 3033 GGACACAGAACTCTGAGTGTAGATGAGCAGGAACTGAGGGCCCTTGGAGCAGGAGAA 3092
 3456 ACCTCAGCTGCTCATCTTTCAGCAGACGGAACCAAGTGTGATGACCCCTACTTGAGGGCT 3515
 3093 GCCTCAGCTGCTCATCTTTCAGCAGACGGAACCAAGTGTGATGACCCCTATTTAGGGCT 3152
 3516 GGATACGGGAGAGATGAGTAAATAGCACATCTCTCACTCAAGCGGAGAGACTCGGATCAGA 3575
 3153 GGATACGGGAGAGATGAGTAAATAGCACATCTCTCACTCAAGCGGAGAGACTCGGATCAGA 3212
 3576 GAGGGCGCTTCCATGTTGCTCAGGCTAATCTCAACTCTGACGCTGAGCAGATCTT 3635
 3213 GAGGGCGCTTCCATGTTGCTCAGGCTAATCTCAACTCTGACGCTGAGCAGATCTT 3272
 3636 CCCAATTCCTCAGATGTCAGAGGAAAGCTCCCGAGAGTAGTACCGGTGGAACCTCTTCTG 3695
 3273 CCCAATTCCTCAGATGTCAGAGGAAAGCTCCCGAGAGTAGTACCGGTGGAACCTCTTCTG 3332
 3696 CGTGCCTTCTCTGCTCTCAAGGGGACCTGCATACGAAGCCTTTTGGGGACTGACGATGA 3755
 3333 CGTGCCTTCTCTGCTCTCAAGGGGACCTGCATACGAAGCCTTTTGGGGACTGACGATGA 3392
 3756 CTTCTGGGGCCCAAGGGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 3815
 3393 CTTCTGGGGCCCAAGGGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 3452
 3816 CCGAGTTCACCTTCCCTGAGCTGCTCTACCGCTGGCCCAACACGCGTCTCTGCTTGT 3875
 3453 CCGAGTTCACCTTCCCTGAGCTGCTCTACCGCTGGCCCAACACGCGTCTCTGCTTGT 3512
 3876 GATGAGAGAGCGGTGACCGTGTGAGATTTGAATTTCTGTGTGGGACCAAGTCTTGGGTGA 3935
 3513 GGTGAGAGAGCGGTGACCGTGTGAGATTTGAATTTCTGTGTGGGACCAAGTCTTGGGTGA 3572
 3936 GATCAACCAAGCAGCAGCTGGATGGTGGCAGGGCTCTGCTGGACATCAAGGCTGAGCC 3995
 3573 GATCAACCAAGCAGCAGCTGGATGGTGGCAGGGCTCTGCTGGACATCAAGGCTGAGCC 3632
 3996 TGGAGCTGTGAGAGCTGACCTCCCTCACTTTGTGCTCTTCAAGGGGCGCATGTGGA 4055
 3633 TGGAGCTGTGAGAGCTGACCTCCCTCACTTTGTGCTCTTCAAGGGGCGCATGTGGA 3692
 4056 CACATCCCTGTTCCAAATGGCCCACTTTTAAAGAGAGGGGATGCTCTCTTGGAGAACCCAGC 4115

Db	3693	CACATCCCTGTTCCAAAGTGCCCACTTTAAAGAGGAGGGATGCTCTCGAGAAGCCAGC	3752
Qy	4116	CAGGGTGGAGCTGCATCACATAGTTCTGGAAAAACCCACAGCTTCTCCCCCTTGGAGAGTCCT	4175
Db	3753	CAGGGTGGAGCTGCATCACATAGTTCTGGAAAAACCCACAGCTTTCCTCCCTTGGAGATCCT	3812
Qy	4176	CTTGAAATGATCCATTAATCCCTCGGCTTCATTTCCTCCGTCACCTCTGTGTGTGTTGTTTA	4235
Db	3813	CTTGAAATGATCCATTAATCCCTCGGCTTCATTTCCTCCGTCACCTCTGTGTGTGTTGTTTA	3872
Qy	4236	CCACCGGTCATCCTTGAGGAAGTCACTTCCACCTTACCTGATCCCAAGTCACTGCTC	4295
Db	3873	CCACCGGTCATCCTTGAGGAAGTCACTTCCACCTTACCTGATCCCAAGTCACTGCTC	3932
Qy	4296	CATTTCGGAAGGAATCGAGCTCTGCTATCGAAGCCCTGGAGAGAAGACAGCTGTCTTCGGA	4355
Db	3933	CATTTCGGAAGGAATCGAGCTCTGCTATCGAAGCCCTGGAGAGAAGACAGCTGTCTTCGGA	3992
Qy	4356	GTTCTAGTTGGCCACTTTGGGATCAGGGATCAGGCTCAAGTGAAGACAAGAAAAATGA	4415
Db	3993	GTTCTAGTTGGCCACTTTGGGATCAGGGATCAGGCTCAAGTGAAGACAAGAAAAATGA	4052
Qy	4416	GACTCTGGTGGGAGGCCCTTGGTGAAAACAGAGATCTCATGCCTGCACACTACTCTGAT	4475
Db	4053	GACTCTGGTGGGAGGCCCTTGGTGAAAACAGAGATCTCATGCCTGCACACTACTCTGAT	4112
Qy	4476	CCCTCCAGCCCGCATAGCCGTACCTTCACCTCTGGATGCCCGCCGAGTTGCTGCACCTTTGT	4535
Db	4113	CCCTCCAGCCCTGCATAGCCGTACCTTCACCTCTGGATGCCCGCCGAGTTGCTGCACCTTTGT	4172
Qy	4536	GGACCACTATTCGAGAGCAGCTGATAGCCCGAGTGACATCGGTGGAGGTTGCTTTGGACAA	4595
Db	4173	GGACCACTATTCGAGAGCAGCTGATAGCCCGAGTGACATCGGTGGAGGTTGCTTTGGACAA	4232
Qy	4596	ACTGCATGGACAGGTGCTGAGCCAGGAGCAGTACGAGAGGGTGCTGGCTGAGAACACGAG	4655
Db	4233	ACTGCATGGACAGGTGCTGAGCCAGGAGCAGTACGAGAGGGTGCTGGCTGAGAACACGAG	4292
Qy	4656	GCCCAGCCAGATCGGAAGCTGTTACGCTTGAGCCAGTCCTGGAGCCGGAAGTGCAAGA	4715
Db	4293	GCCCAGCCAGATCGGAAGCTGTTACGCTTGAGCCAGTCCTGGAGCCGGAAGTGCAAGA	4352
Qy	4716	TGGACTCTACAAGCCCTGAAGGAGACCCATCTCTCACCTCATTTATGGAACCTCTGGGAGAA	4775
Db	4353	TGGACTCTACAAGCCCTGAAGGAGACCCATCTCTCACCTCATTTATGGAACCTCTGGGAGAA	4412
Qy	4776	GGGAGCAAAAAGGGAATCTGCGCACTCAGCAGCTGAAGTATCAACACAGCCCTTGACC	4835
Db	4413	GGGAGCAAAAAGGGAATCTGCGCACTCAGCAGCTGAAGTATGAACACAGCCCTTGACC	4472
Qy	4836	CTTGAGTCTCGGCTTTGGCTGACCCCTTTGGGTCTCAGTTTCTTCTCTGCAAAACAAG	4895
Db	4473	CTTGAGTCTCGGCTTTGGCTGACCCCTTTGGGTCTCAGTTTCTTCTCTGCAAAACAAG	4532
Qy	4896	TTGCCATCTGGTTTGCCCTTCAGACACTAAAGTAATGGAACTTTTGATGATGCCCTTTGCTGG	4955
Db	4533	TTGCCATCTGGTTTGCCCTTCAGACACTAAAGTAATGGAACTTTTGATGATGCCCTTTGCTGG	4592
Qy	4956	GCATTATGTGTCATGTCAGGGATGCCACAGGGGGGCCCGAGTCAGGTGCGCTTAACAGCA	5015
Db	4593	GCATTATGTGTCATGTCAGGGATGCCACAGGGGGGCCCGAGTCAGGTGCGCTTAACAGCA	4652
Qy	5016	TCTCAGGGAATGTCATCTGGAGCTGCAGACCCCTCGACACTCATAGAGCTCATCT	5075
Db	4653	TCTCAGGGAATGTCATCTGGAGCTGCAGACCCCTCGACACTCATAGAGCTCATCT	4712
Qy	5076	GGTGGCCACAGCAGCCAAAGCCTAGAGCCCTCCGGATCCCATCCAGGCGCAAGAGGAATA	5135
Db	4713	GGTGGCCACAGCAGCCAAAGCCTAGAGCCCTCCGGATCCCATCCAGGCGCAAGAGGAATA	4772
Qy	5136	GGAGGGACATGGAAACCATTTGCTCTGGCTGTGTGCAGAGGGTGAGCCCCCAAAATGGGGT	5195

Db	4773	GGAGGGACATGGAAACAATTTGGCTCTGGCTGTGTCTACAGGGTGAGCCCCAAAATTTGGGGT	4833
Qy	5196	TCACGTGGGAGGCCACGTCGGATTCTTTGGCTTTGTACAGGAAGATCTACAAAGAGCAAGCC	5255
Db	4833	TCAGCGTGGGAGGCCACGTCGATTCTTTGGCTTTGTACAGGAAGATCTACAAAGAGCAAGCC	4892
Qy	5256	AACAGAGTAAGTGGAAAGGAAGTTATTACAGAAAATAAAGGAGTATCACAGCTCTTTTAG	5315
Db	4893	AACAGAGTAAGTGGAAAGGAAGTTATTACAGAAAATAAAGGAGTATCACCTGCTCTTTTAG	4952
Qy	5316	AAATTGTCTAGCAGGCTTTCCAGTGTTTTACCAGAAAACCCCTATAAATTAAAAATTTTTT	5375
Db	4953	AAATTGTCTAGCAGACCTTCCAGTGTTTTACCAGAAAACCCCTATAAATTAAAAATTTTTT	5012
Qy	5376	ACTTAATTTAAGAAATTAATAAAAAATACAAAAAGAAAAATGAAAAATAAAGAAATAAGAA	5435
Db	5013	ACTTAATTTAAGAAATTAATAAAAAATACAAAAAGAAAAATGAAAAATAAAGAAATAAGAA	5072
Qy	5436	GTTA 5439	
Db	5073	GTTA 5076	
RESULT 5			
AAD12951			
ID	AAD12951 standard; cDNA; 5122 BP.		
XX	AAD12951;		
XX	16-OCT-2001 (first entry)		
XX	Human G-protein coupled receptor-8 (GCRC-8) cDNA.		
XX	Human; G-protein coupled receptor-8; GCRC-8; cytosolic; hepatotropic;		
KW	virucide; antiinflammatory; anticonvulsant; antiemetic; neuroprotective;		
KW	nootropic; cerebroprotective; hypotensive; tranquiliser; vulnerary;		
KW	ophthalmological; cell proliferative disorder; actinic keratosis;		
KW	anorectic; arteriosclerosis; atherosclerosis; cirrhosis; hepatitis;		
KW	psoriasis; cancer; neurological disorder; stroke; Alzheimer's disease;		
KW	Huntington's disease; Parkinson's disease; cardiovascular disorder;		
KW	epilepsy; hypertension; varicose vein; vasculitis; dysphagia; dyspepsia;		
KW	anorexia; gastrointestinal disorder; pancreatitis; autoimmune disorder;		
KW	Addison's disease; Crohn's disease; acquired immune deficiency syndrome;		
KW	AIDS; uveitis; infection; trauma; metabolic disorder; diabetes; obesity;		
XX	osteoporosis; transgenic animal; gene therapy; ss.		
XX	Homo sapiens.		
OS	Homo sapiens.		
XX	Key Location/Qualifiers		
FH	CDS 489..4910		
FT	/*tag= a		
FT	/product= "Human GCRC-8 protein"		
XX	WO200157085-A2.		
XX	09-AUG-2001.		
XX	01-FEB-2001; 2001WO-US03455.		
XX	02-FEB-2000; 2000US-0180093.		
PR	11-FEB-2000; 2000US-0182045.		
XX	(INCY-) INCYTE GENOMICS INC.		
XX	Baughn MR, Au-Young J, Yue H;		
XX	WPI; 2001-488869/53.		
DR	P-PSDB; AAE06758.		
XX	Novel isolated human G-protein coupled receptor useful for diagnosing,		
FT	preventing and treating cell proliferative, neurological,		
FT	cardiovascular, gastrointestinal, autoimmune/inflammatory and metabolic		
PT	disorders -		

Claim 5; Page 131-132; 138pp; English.
 The present sequence is human G-protein coupled receptor-8 (GCR8-8) cDNA. The present invention relates to GCR8 protein and nucleic acids encoding them. GCR8 protein, its agonist or antagonist are useful for treating diseases or conditions associated with decreased expression or overexpression of functional GCR8 in a patient, where the disorder is selected from cell proliferative disorders such as actinic keratosis, arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, psoriasis, and cancer, neurological disorders such as epilepsy, stroke, Alzheimer's disease, Huntington's disease, Parkinson's disease, cardiovascular diseases such as hypertension, vasculitis, varicose veins, gastrointestinal disorders such as dysphagia, dyspepsia, anorexia, nausea, pancreatitis, autoimmune/inflammatory disorders such as acquired immunodeficiency syndrome (AIDS), Addison's disease, Crohn's disease, urethritis, viral, bacterial, fungal, parasitic, protozoal, helminthic infections, trauma and metabolic disorders such as diabetes, obesity, osteoporosis. GCR8 proteins and their cDNAs are used to assess the effects of exogenous compounds on the expression of GCR8 sequences. GCR8 cDNA is useful to create knock in humanised animals (pigs) or transgenic animals (mice or rats) to model human disease, for therapeutic or diagnostic purposes, for somatic or germline gene therapy, to generate hybridisation probes useful in mapping the naturally occurring genomic sequence, and in molecular biological techniques.

Query Match 87.5%; Score 4765.6; DB 22; Length 5122;
 Best Local Similarity 96.9%; Pred. No. 0;
 Matches 4933; Conservative 0; Mismatches 24; Indels 133; Gaps 2;

XX	35	GAGCCAGACGCGGGGCTCCACTGCGGTCTGAAAGCCCATTCCTGCTCTGGGCTC	94	CGGTGAGACACCCGCTCAGCCAGAGAAAGACGAGTGGCATGGAGTGGCTCGTACCTGG	694
PS	1	GAGCCAGACGCGGGGCTCCACTGCGGTCTGAAAGCCCATTCCTGCTCTGGGCTC	60	CGGTGAGACACCCGCTCAGCCAGAGAAAGACGAGTGGCATGGAGTGGCTCGTACCTGG	660
CC	95	CTCCACACCATCTCTCAGCTTTCAGCTCAAGGGTTGATCTCAGGAGTCCAGGACC	154	TGGCTCAGTATGGGAGCAGCGGCTGGGACCTTAGCCCTCCATACCTGGGAGCAGATGG	754
CC	61	CTCCACACCATCTCTCAGCTTTCAGCTCAAGGGTTGATCTCAGGAGTCCAGGACC	120	TGGCTCAGTATGGGAGCAGCGGCTGGGACCTTAGCCCTCCATACCTGGGAGCAGATGG	720
CC	155	CAGGAGAGGAGAAATCTGAGGAACAACAGACAGTGGGCTTGCCACACCCCATCTCCC	214	GGCTGAGTCACTGTGCGCCCAAGCCAGGAAGGGCAGGCCACTCTCCCTCATTTCCCT	814
CC	121	CAGGAGAGGAGAAATCTGAGGAACAACAGACAGTGGGCTTGCCACACCCCATCTCCC	180	GGCTGAGTCACTGTGCGCCCAAGCCAGGAAGGGCAGGCCACTCTCCCTCATTTCCCT	780
CC	215	GTACACACATCTCCCTCAGCTCACCCTCCCTGCTGCGCTTGGACCCCATCCAGGAC	274	ACAGCCCAAGTGAACCCCACTGCGGTCTCCAGCCAAACCCCACTCCACCCAGTGTAA	840
CC	181	GTACACACATCTCCCTCAGCTCACCCTCCCTGCTGCGCTTGGACCCCATCCAGGAC	240	TGCTCTGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	934
CC	275	CTCCCTATCAGCTGACTCTTCCAGTGTCTGAGGCGCCCTCTGGGCTCTCTCCCTCC	334	TGCTCTGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	900
CC	241	CTCCCTATCAGCTGACTCTTCCAGTGTCTGAGGCGCCCTCTGGGCTCTCTCCCTCC	300	ACAGCCCAAGTGAACCCCACTGCGGTCTCCAGCCAAACCCCACTCCACCCAGTGTAA	840
CC	335	GGCTTTTCTTACCATCTCCCTCTATCGGCTCTATCTGAGTGGCTTGGGATTTATAA	394	TGCTCTGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	934
CC	301	GGCTTTTCTTACCATCTCCCTCTATCGGCTCTATCTGAGTGGCTTGGGATTTATAA	360	TGCTCTGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	900
CC	395	AAGTGGTTCGATGCTGAAATGAGAGCGTGAAGAGCCGAGCCAGGAGAGAGCTGTT	454	ACAGCCCAAGTGAACCCCACTGCGGTCTCCAGCCAAACCCCACTCCACCCAGTGTAA	840
CC	361	AAGTGGTTCGATGCTGAAATGAGAGCGTGAAGAGCCGAGCCAGGAGAGAGCTGTT	420	TGCTCTGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	934
CC	455	CTCTGCTGCTGATACCTCTCACCCTCGGAGACATCCCCAGACACCTCTTTACTCCG	514	TGCTCTGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	900
CC	421	CTCTGCTGCTGATACCTCTCACCCTCGGAGACATCCCCAGACACCTCTTTACTCCG	480	ACAGCCCAAGTGAACCCCACTGCGGTCTCCAGCCAAACCCCACTCCACCCAGTGTAA	840
CC	515	GGAAGAGATGGTGGCGGAGCTGGGCGGCTGCTGCTGTTACTTGGAGTTCTTGAAGA	574	TGCTCTGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	934
CC	481	GGACAGAGATGGTGGCGGAGCTGGGCGGCTGCTGCTGTTACTTGGAGTTCTTGAAGA	540	TGCTCTGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	900
CC	575	AGAGAGAGCTGAAGGAGTTTCAGCTTCTGCTGCGCAATGAAGGCGACTCCAGAGCT	634	ACAGCCCAAGTGAACCCCACTGCGGTCTCCAGCCAAACCCCACTCCACCCAGTGTAA	840
CC	541	AGAGAGAGCTGAAGGAGTTTCAGCTTCTGCTGCGCAATGAAGGCGACTCCAGAGCT	600	TGCTCTGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	934

QY	635	CGGTGAGACACCCGCTCAGCCAGAGAAAGACGAGTGGCATGGAGTGGCTCGTACCTGG	694
DB	601	CGGTGAGACACCCGCTCAGCCAGAGAAAGACGAGTGGCATGGAGTGGCTCGTACCTGG	660
QY	695	TGGCTCAGTATGGGAGCAGCGGCTGGGACCTTAGCCCTCCATACCTGGGAGCAGATGG	754
DB	661	TGGCTCAGTATGGGAGCAGCGGCTGGGACCTTAGCCCTCCATACCTGGGAGCAGATGG	720
QY	755	GGCTGAGTCACTGTGCGCCCAAGCCAGGAAGGGCAGGCCACTCTCCCTCATTTCCCT	814
DB	721	GGCTGAGTCACTGTGCGCCCAAGCCAGGAAGGGCAGGCCACTCTCCCTCATTTCCCT	780
QY	815	ACAGCCCAAGTGAACCCCACTGCGGTCTCCAGCCAAACCCCACTCCACCCAGTGTAA	874
DB	781	ACAGCCCAAGTGAACCCCACTGCGGTCTCCAGCCAAACCCCACTCCACCCAGTGTAA	840
QY	875	TGCTCTGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	934
DB	841	TGCTCTGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	900
QY	935	GACAGTCTGCTGACACATCTGCGCGCTCGAGAGAAATCTCTGCTCTACCTCTCTACC	994
DB	901	GACAGTCTGCTGACACATCTGCGCGCTCGAGAGAAATCTCTGCTCTACACGCTCTACC	960
QY	995	AAGCTCTTCCAAAGCTCCCGACACCATGATCTCCAAGCCAGGAGTCAACCCACGCCCA	1054
DB	961	AAGCTCTTCCAAAGCTCCCGACACCATGATCTCCAAGCCAGGAGTCAACCCACGCCCA	1020
QY	1055	CATCCACAGCAGTGTGGGAGCTGGGGATCCCACTCAGCCAGCCCTAGCACCAGAG	1114
DB	1021	CATCCACAGCAGTGTGGGAGCTGGGGATCCCACTCAGCCAGCCCTAGCACCAGAG	1080
QY	1115	AGCAGAGGCTCTCTGGGACCCCAATGGGCTCTGGATGAAACCTCAGGAATTTTACTACACAG	1174
DB	1081	AGCAGAGGCTCTCTGGGACCCCAATGGGCTCTGGATGAAACCTCAGGAATTTTACTACACAG	1140
QY	1175	AAATCAGAGAAAGAGAGAGAGAGAAATCAGAGAAAGGAGGCGCCCATGGGAGCGGTGG	1234
DB	1141	AAATCAGAGAAAGAGAGAGAGAGAAATCAGAGAAAGGAGGCGCCCATGGGAGCGGTGG	1200
QY	1235	TAGGAAACCCCGCCACAGCGGCACACAGCCTACAGCCCAACCCATGGGAGCGCTT	1294
DB	1201	TAGGAAACCCCGCCACAGCGGCACACAGCCTACAGCCCAACCCATGGGAGCGCTT	1260
QY	1295	CTGTGAGAGAGAGCTCTGTTTCCACATGGCCCTGGAATAATGAGGATTTTAAACCAAAAT	1354
DB	1261	CTGTGAGAGAGAGCTCTGTTTCCACATGGCCCTGGAATAATGAGGATTTTAAACCAAAAT	1320
QY	1355	TCACACAGCTGCTACTTCTCAAAAGACCTCACCCAGAGCCAAAGATCCCTGGTCAAGA	1414
DB	1321	TCACACAGCTGCTACTTCTCAAAAGACCTCACCCAGAGCCAAAGATCCCTGGTCAAGA	1380
QY	1415	GAGCTGCGCTGATTTATGAGGAGAGAAATCAGAGACATTTTAAATTTGAGATTCAGAGCTTAT	1474
DB	1381	GAGCTGCGCTGATTTATGAGGAGAGAAATCAGAGACATTTTAAATTTGAGATTCAGAGCTTAT	1440
QY	1475	TTGGCCAGAGCTGGATACCAAGAACTCCGATAGTATCATCTGAGGAGGCTGCTGGA	1534
DB	1441	TTGGCCAGAGCTGGATACCAAGAACTCCGATAGTATCATCTGAGGAGGCTGCTGGA	1500
QY	1535	TTGGGAGTCAACACTGGCCAGGAGGAGTGAAGGAGCCCTGGGGAGAGAGCCAGCTGTATG	1594
DB	1501	TTGGGAGTCAACACTGGCCAGGAGGAGTGAAGGAGCCCTGGGGAGAGAGCCAGCTGTATG	1560
QY	1595	GGGACCGCTTCAGAGTCTTCTTACTTTCAGTGTGAGAGCTGGCCAGCTTCAAGGTGG	1654
DB	1561	GGGACCGCTTCAGAGTCTTCTTACTTTCAGTGTGAGAGCTGGCCAGCTTCAAGGTGG	1620
QY	1655	TGAGTCTCGCTGAGCTCATCGGAAAGATGGAGAGCCACTCCGCTCCCATTTACACAGA	1714
DB	1621	TGAGTCTCGCTGAGCTCATCGGAAAGATGGAGAGCCACTCCGCTCCCATTTACACAGA	1680
QY	1715	TCCTGCTAGGCGAGAGCGGCTGCTTCTTCTCATCTCTGATGATGATGAGCAGGATGGG	1774

Db 1681 TCTCTGCTAGGCGAGCGGCTGCTCTTCACTCCGATGGTGTAGATGAGCCAGATGGG 1740
Qy 1775 TCTTCCAGAGCCGAGTCTGAGCTCTGTCTGCACTGGAGCCAGCCACAGCGGGGATG 1834
Db 1741 TCTTGCAGAGCCGAGTCTGAGCTCTGTCTGCACTGGAGCCAGCCACAGCGGGGATG 1800
Qy 1835 CACTGCTGGGAGTTTGTCTGGGAAAACTATACCTTCCCGAGGCACTCTTCTGTATCACGG 1894
Db 1801 CACTGCTGGGAGTTTGTCTGGGAAAACTATACCTTCCCGAGGCACTCTTCTGTATCACGG 1860
Qy 1895 CTGGACACAGCTCTGAGAAACCTCATCTCTTCTTGGAGGACGACGTTGGGTAGAGG 1954
Db 1861 CTGGACACAGCTCTGAGAAACCTCATCTCTTCTTGGAGGACGACGTTGGGTAGAGG 1920
Qy 1955 TCTCGGGTCTCTCAGTCCAGCAGAAAGGAATATTTCTACAGATATTTTACAGATGAAA 2014
Db 1921 TCTCGGGTCTCTCAGTCCAGCAGAAAGGAATATTTCTACAGATATTTTACAGATGAAA 1980
Qy 2015 GGCAAGCAATTAGAGCCTTTAGTTGGTCAAAATCAAAACAAAGAGCTCTGGGCCCTGTGTC 2074
Db 1981 GGCAAGCAATTAGAGCCTTTAGTTGGTCAAAATCAAAACAAAGAGCTCTGGGCCCTGTGTC 2040
Qy 2075 TTGTGCTTGGGTCTGCTGGCTGGCTGCACTTGGCTGATGAGCAGATGAAGCGGAAGG 2134
Db 2041 TTGTGCTTGGGTCTGCTGGCTGGCTGCACTTGGCTGATGAGCAGATGAAGCGGAAGG 2100
Qy 2135 AAAAATCACACTGACTTCCAGAGACACCAACAACTCTGTCTCAATTACCTTTGCCCAGG 2194
Db 2101 AAAAATCACACTGACTTCCAGAGACACCAACAACTCTGTCTCAATTACCTTTGCCCAGG 2160
Qy 2195 CTCTCAAGCTCAGGCATTTGGGACCCAGCTCAGAGACCTCTGCTCTCTGGCTGTGAGG 2254
Db 2161 CTCTCAAGCTCAGGCATTTGGGACCCAGCTCAGAGACCTCTGCTCTCTGGCTGTGAGG 2220
Qy 2255 GCATCTGGCAAAAAAGACCTTTTCAATGTCAGATGACCTCAGGAAGCATGGGTAGATG 2314
Db 2221 GCATCTGGCAAAAAAGACCTTTTCAATGTCAGATGACCTCAGGAAGCATGGGTAGATG 2280
Qy 2315 GGGCCATCATCTCCACCTCTTGAAGATGGGTATTTCTTCAAGAGCACCCCATCTCTGA 2374
Db 2281 GGGCCATCATCTCCACCTCTTGAAGATGGGTATTTCTTCAAGAGCACCCCATCTCTGA 2340
Qy 2375 GCTACAGCTTCAATCACTCTGTTTCCAAAGAGTCTTTGACAGCAATGTCTATGCTTGG 2434
Db 2341 GCTACAGCTTCAATCACTCTGTTTCCAAAGAGTCTTTGACAGCAATGTCTATGCTTGG 2400
Qy 2435 AGGATGAGAAAGGGAGAGGTAAACATTTCTAATGTCATCATAGATTTGGAAAAAGACGTAG 2494
Db 2401 AGGATGAGAAAGGGAGAGGTAAACATTTCTAATGTCATCATAGATTTGGAAAAAGACGTAG 2460
Qy 2495 AAGCATATGGAATACATGGCCCTGTTGGGGCATCAACACACAGTTTCTTATTTGGGCTGT 2554
Db 2461 AAGCATATGGAATACATGGCCCTGTTGGGGCATCAACACACAGTTTCTTATTTGGGCTGT 2520
Qy 2555 TAAGTGATGAGGGGAGAGAGATGGAGAACATCTTTCACCTCCGGCTGTCTCAGGGGA 2614
Db 2521 TAAGTGATGAGGGGAGAGAGATGGAGAACATCTTTCACCTCCGGCTGTCTCAGGGGA 2580
Qy 2615 GGAACTGATGAGTGGGTCCCCTGCCAGTGTCTGTCAGCCACACTCTCTGGAGT 2674
Db 2581 GGAACTGATGAGTGGGTCCCCTGCCAGTGTCTGTCAGCCACACTCTCTGGAGT 2640
Qy 2675 CCCTCCAGCTTGTACAGAGATCTGGAAACAAACGTTCTGTACACAGTGTAGTGGCCCAT 2734
Db 2641 CCCTCCAGCTTGTATGAGATCTGGAAACAAACGTTCTGTACACAGTGTAGTGGCCCAT 2700
Qy 2735 TCGAAGAAATGGGCACTGTGTAGAAACAGACATGGAGCTTTAGTGTGCATTTCTGCA 2794
Db 2701 TCGAAGAAATGGGCACTGTGTAGAAACAGACATGGAGCTTTAGTGTGCATTTCTGCA 2760
Qy 2795 TTTAAATTCAGCCGCCACGTGAAGAAAGCTTTCAGTGTATTTGAGGCGAGCAGCATCAA 2854

Db 2761 TTTAAATTCAGCCGCCACGTGAAGAAAGCTTTCAGTGTATTTGAGGCGAGCAGCAGCATCAA 2820
Qy 2855 CATGGAGCCCCACCATGCTAGTCTCTGTTTCAAGTGGTCCCCAGTCAAGATCCTTATGGC 2914
Db 2821 CATGGAGCCCCACCATGCTAGTCTCTGTTTCAAGTGGTCCCCAGTCAAGATCCTTATGGC 2880
Qy 2915 AGATTCTTCTTCTCGTCTCAAGGTCAACAGAAACCTGAAAGAGCTGACCTTAAGTGAA 2974
Db 2881 AGATTCTTCTTCTCGTCTCAAGGTCAACAGAAACCTGAAAGAGCTGACCTTAAGTGAA 2940
Qy 2975 ACTCGCTGAGCACTCTGCACTGAAGAGTCTTTTGAAGCCCTGAGACGCTCTGCTGCC 3034
Db 2941 ACTCGCTGAGCACTCTGCACTGAAGAGTCTTTTGAAGCCCTGAGACGCTCTGCTGCC 3000
Qy 3035 TCTTGGAGACCTCTCGGTTGGCTGTGGCCTCACAGCTGAGGACTGCAAGGACCTTG 3094
Db 3001 TCTTGGAGACCTCTCGGTTGGCTGTGGCCTCACAGCTGAGGACTGCAAGGACCTTG 3060
Qy 3095 CTTTGGGCTGAGAGCCAAACAGACCTTGAACGAGCTGGACCTGAGCTTCAATGTGCTCA 3154
Db 3061 CTTTGGGCTGAGAGCCAAACAGACCTTGAACGAGCTGGACCTGAGCTTCAATGTGCTCA 3120
Qy 3155 CGGATGCTGGAGCCAAACACCTTTGCCAGAGACTGAGACGCGAGCTGCAAGCTACAGC 3214
Db 3121 CGGATGCTGGAGCCAAACACCTTTGCCAGAGACTGAGACGCGAGCTGCAAGCTACAGC 3180
Qy 3215 GACTGCACTGCTGAGCTGTGGCCTCACGCTGCTGCTGTCAGGACCTGGCCTCTGTGC 3274
Db 3181 GACTGCACTGCTGAGCTGTGGCCTCACGCTGCTGCTGCCAGGACCTGGCCTCTGTGC 3240
Qy 3275 TTAGTGCAGCCCGACGCTGAAGAGCTAGACCTTGCACAGAAACAACTGATGACGTTG 3334
Db 3241 TTAGTGCAGCCCGACGCTGAAGAGCTAGACCTTGCACAGAAACAACTGATGACGTTG 3300
Qy 3335 GCCTGCACTGCTCTGTCAGGGGCTCAGGCATCTCTGCTGCANACTCATAGCCTGGGGC 3394
Db 3301 GCCTGCACTGCTCTGTCAGGGGCTCAGGCATCTCTGCTGCANACTCATAGCCTGGGGC 3360
Qy 3395 TGGACACACAACTCTGATGATGATGAGTGAAGGACAGGAACTGAGGGCCCTGGAGCAGGAGA 3454
Db 3361 TGGACACACAACTCTGATGATGATGATGAGTGAAGGACAGGAACTGAGGGCCCTGGAGCAGGAGA 3420
Qy 3455 AACCTCAGCTCTCATCTTCAGCAGACGGAACCAAGTGTGATGACCCCTACTGAGGGCC 3514
Db 3421 AACCTCAGCTCTCATCTTCAGCAGACGGAACCAAGTGTGATGACCCCTACTGAGGGCC 3480
Qy 3515 TGGATACGGGAGATGATGATATAGCACATCTCTCACTCAAGGGCAGAGACTCGCATCAG 3574
Db 3481 TGGATACGGGAGATGATGATATAGCACATCTCTCACTCAAGGGCAGAGACTCGCATCAG 3540
Qy 3575 AGAGGGGCTTCCCATGTTGCTCAGGCTAATCTCAAACTCTCTGACGCTGACAAAGATCT 3634
Db 3541 AGAGGGGCTTCCCATGTTGCTCAGGCTAATCTCAAACTCTCTGACGCTGACAAAGATCT 3600
Qy 3635 TCCCAATTTGCTGAGATTGACAGGAAAGCTCCCAAGAGTGTACCCGTTGGAACTCTTGT 3694
Db 3601 TCCCAATTTGCTGAGATTGACAGGAAAGCTCCCAAGAGTGTACCCGTTGGAACTCTTGT 3660
Qy 3695 GGTGCTCTTCTGCTCTCAAGGGGACCTGATACGAAGCCTTTGGGGACTGCGATG 3754
Db 3661 GGTGCTCTTCTGCTCTCAAGGGGACCTGATACGAAGCCTTTGGGGACTGCGATG 3720
Qy 3755 ACTTCTGGGGCCCCACGGGCTGTGGCTACTGAGTGTAGTGTGACAAAGAAAGAACTTGT 3814
Db 3721 ACTTCTGGGGCCCCACGGGCTGTGGCTACTGAGTGTAGTGTGACAAAGAAAGAACTTGT 3780
Qy 3815 ACCGAGTTCACTTCCCTGTAGTGGCTCTACCGCTGGCCCCAACACGGGTCTCTGCTTGT 3874
Db 3781 ACCGAGTTCACTTCCCTGTAGTGGCTCTACCGCTGGCCCCAACACGGGTCTCTGCTTGT 3840
Qy 3875 TGAATGAGAAAGCGGTGACCGTTGAGATTTGAATTTCTGTGTGGGACCAAGTTCTCTGGGTG 3934
Db 3841 TGAATGAGAAAGCGGTGACCGTTGAGATTTGAATTTCTGTGTGGGACCAAGTTCTCTGGGTG 3900

CC NAC proteins are used for screening modulators that modulates apoptosis,
 CC cytokine production, cytokine receptor signalling and other cellular
 CC processes. NAC can act as an immunogen for the production of polyclonal
 CC and monoclonal antibodies. It can also be used to diagnose and treat
 CC inflammatory disorders such as sepsis, fibrosis and arthritis and cancer
 CC pathologies such as adenocarcinomas and leukaemias.
 XX

SQ Sequence 4422 BP; 1043 A; 1224 C; 1240 G; 915 T; 0 other;

Query Match		76.2%;	Score 4148;	DB 22;	Length 4422;
Best Local Similarity		97.0%;	Pred. No. 0;		
Matches 4290;		Conservative 0;	Mismatches 0;	Indels 132;	Gaps 1;
QY	523	ATGGCTGGCGGAGCCCTGGGGCCGCTGGCTGTTACTTGGAGTTCTGAAGAGGAGGAG	582		
DB	1	ATGGCTGGCGGAGCCCTGGGGCCGCTGGCTGTTACTTGGAGTTCTGAAGAGGAGGAG	60		
QY	583	CTGAAGGAGTTCAGCTTCTGCTCGCAATAAAGCGCACTCCAGGAGCTTTCGGGTGAG	642		
DB	61	CTGAAGGAGTTCAGCTTCTGCTCGCAATAAAGCGCACTCCAGGAGCTTTCGGGTGAG	120		
QY	643	ACACCGCTCAGCCAGAGAGAGAGTGGCATGGAGTGGCCCTCGTACCTGTGGTCTCAG	702		
DB	121	ACACCGCTCAGCCAGAGAGAGAGTGGCATGGAGTGGCCCTCGTACCTGTGGTCTCAG	180		
QY	703	TATGGGAGCAGCGGCGCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGCGTGAGG	762		
DB	181	TATGGGAGCAGCGGCGCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGCGTGAGG	240		
QY	763	TCACTGTGGCCCAAGCCAGAGAGGGGAGGCGCACTCTCCCTCAATCCCTCCACAGCCCA	822		
DB	241	TCACTGTGGCCCAAGCCAGAGAGGGGAGGCGCACTCTCCCTCAATCCCTCCACAGCCCA	300		
QY	823	AGTGAACCCCACTGGGGTCTCCAGCAACCCACCTCCACCGCAGTGTAAATGCCCTGG	882		
DB	301	AGTGAACCCCACTGGGGTCTCCAGCAACCCACCTCCACCGCAGTGTAAATGCCCTGG	360		
QY	883	ATCCATGAATTCGGCGGGGTGCACCCAGGGCTCAGAGAGAGGGTTCAGAGACAGCTG	942		
DB	361	ATCCATGAATTCGGCGGGGTGCACCCAGGGCTCAGAGAGAGGGTTCAGAGACAGCTG	420		
QY	943	CTGACACATCTGGAGCGCGCTGGAGAGAAATCTTGCCTCACTCCTCAACAGCTCTT	1002		
DB	421	CTGACACATCTGGAGCGCGCTGGAGAGAAATCTTGCCTCACTCCTCAACAGCTCTT	480		
QY	1003	CCAGCTCCCCAGACCATGAGTCTCAAGCCAGGAGTCAACCAACGCCCCCACTCCACA	1062		
DB	481	CCAAGCTCCCCAGACCATGAGTCTCAAGCCAGGAGTCAACCAACGCCCCCACTCCACA	540		
QY	1063	GCAGTGCTGGGGAGCTGGGGATCCCACTCAGCCAGCCTAGCACCCAGAGAGAGGAG	1122		
DB	541	GCAGTGCTGGGGAGCTGGGGATCCCACTCAGCCAGCCTAGCACCCAGAGAGAGGAG	600		
QY	1123	GCTCCTGGGACCAATGGCTCTGGATGAACCTCAGGAATTTTACACAGAAATCAGA	1182		
DB	601	GCTCCTGGGACCAATGGCTCTGGATGAACCTCAGGAATTTTACACAGAAATCAGA	660		
QY	1183	GAAGAGAGAGAGAGAAATCAGAGAAAGGACGCCCCCATGGGACGGGTGGTAGAACG	1242		
DB	661	GAAGAGAGAGAGAGAAATCAGAGAAAGGACGCCCCCATGGGACGGGTGGTAGAACG	720		
QY	1243	CCCCCAGCGGCACACCGCTTACAGCCCCACACCCACCTAGGGAGCTTCTGTGAGA	1302		
DB	721	CCCCCAGCGGCACACCGCTTACAGCCCCACACCCATAGGGAGCTTCTGTGAGA	780		
QY	1303	GAGAGGCTCTGTTCCACATGGCCCTGGAAAAATGAGGATTTTAAACAAAAATTCACACAG	1362		
DB	781	GAGAGGCTCTGTTCCACATGGCCCTGGAAAAATGAGGATTTTAAACAAAAATTCACACAG	840		
QY	1363	CTGCTACTTCTACAAAGACCTCACCCAGAGGCAAGATCCCTGGTCAAGAGAGCTGG	1422		
DB	841	CTGCTACTTCTACAAAGACCTCACCCAGAGGCAAGATCCCTGGTCAAGAGAGCTGG	900		

QY	1423	CCTGATTATGTGGAGAGAAATCGAGGACATTTAATTGAGATCAGAGACTTATTTGGCCCA	1482
DB	901	CCTGATTATGTGGAGAGAAATCGAGGACATTTAATTGAGATCAGAGACTTATTTGGCCCA	960
QY	1483	GGCCTGGATACCCAGAAACCTCGCATAGTCACTACGAGGGGCTGCTGGAATTCGGAG	1542
DB	961	GGCCTGGATACCCAGAAACCTCGCATAGTCACTACGAGGGGCTGCTGGAATTCGGAG	1020
QY	1543	TCAAACACTGGCCAGGCGAGTGAAGAAAGCCTGGGGAGAGGCCAGCTGTATGGGACCGC	1602
DB	1021	TCAAACACTGGCCAGGCGAGTGAAGAAAGCCTGGGGAGAGGCCAGCTGTATGGGACCGC	1080
QY	1603	TTCCAGCATGTCTTCTTCACTTCACTGAGAGAGCTGGGCCAGTCCAAAGTGTGAGTCTC	1662
DB	1081	TTCCAGCATGTCTTCTTCACTTCACTGAGAGAGCTGGGCCAGTCCAAAGTGTGAGTCTC	1140
QY	1663	GCTGAGCTCATCGGAAAGATGGGACAGCACTCCGGCTCCCATTAGACAGATCCTGTCT	1722
DB	1141	GCTGAGCTCATCGGAAAGATGGGACAGCACTCCGGCTCCCATTAGACAGATCCTGTCT	1200
QY	1723	AGGCCAGAGCGGCTGCTCTTCACTCTCGATGTGTAGATGAGCCAGGATGGGCTCTTGCA	1782
DB	1201	AGGCCAGAGCGGCTGCTCTTCACTCTCGATGTGTAGATGAGCCAGGATGGGCTCTTGCA	1260
QY	1783	GAGCCGAGTCTGAGCTCTGCTGCACTGGAGCCAGCCACAGCCGCGGATGCACTGCTG	1842
DB	1261	GAGCCGAGTCTGAGCTCTGCTGCACTGGAGCCAGCCACAGCCGCGGATGCACTGCTG	1320
QY	1843	GGCAGTTTGTGGGAAACATATACTCCCGAGGCACTCTTCTGATCAGGCTCGGACC	1902
DB	1321	GGCAGTTTGTGGGAAACATATACTCCCGAGGCACTCTTCTGATCAGGCTCGGACC	1380
QY	1903	ACAGCTCTGCAGAACTCACTTCTTCTTGGAGCAGGCACTTGGGTAGAGTCTCTGGGG	1962
DB	1381	ACAGCTCTGCAGAACTCACTTCTTCTTGGAGCAGGCACTTGGGTAGAGTCTCTGGGG	1440
QY	1963	TTCTCTGAGTCAGAGGAGAAATATTTCACAGATTTTCAAGATGAAGGCAAGCA	2022
DB	1441	TTCTCTGAGTCAGAGGAGAAATATTTCACAGATTTTCAAGATGAAGGCAAGCA	1500
QY	2023	ATTAGAGCCTTTAGGTTGGTCAAATCAAACAAGAGCTCTGGGCCCTGTCTTGTGCC	2082
DB	1501	ATTAGAGCCTTTAGGTTGGTCAAATCAAACAAGAGCTCTGGGCCCTGTCTTGTGCC	1560
QY	2083	TGGGTGCTCTGCTGCTGCTGCACTTGCCTGTAGTCAGCAGATGAAGCGGAAGAAACTC	2142
DB	1561	TGGGTGCTCTGCTGCTGCTGCACTTGCCTGTAGTCAGCAGATGAAGCGGAAGAAACTC	1620
QY	2143	ACACTGACTTCCAAAGACCAACCAACCCCTCTGTCTACATTTACCTTGGCCAGGCTCTC	2202
DB	1621	ACACTGACTTCCAAAGACCAACCAACCCCTCTGTCTACATTTACCTTGGCCAGGCTCTC	1680
QY	2203	GCTCAGCATTTGGGACCCAGCTCAGAGACCTCTGCTCTCTGCTGCTGAGGGCATCTGG	2262
DB	1681	GCTCAGCATTTGGGACCCAGCTCAGAGACCTCTGCTCTCTGCTGCTGAGGGCATCTGG	1740
QY	2263	CAAAAAAGACCCCTTTTCACTGCTCAGATGACCTCAGGAGCATGGGTAGATGGGGCATC	2322
DB	1741	CAAAAAAGACCCCTTTTCACTGCTCAGATGACCTCAGGAGCATGGGTAGATGGGGCATC	1800
QY	2323	ATCTCCACCTTCTTGAAGATGGGTATTTCTTCAAGAGCACCCCATCCCTCTGAGCTACAG	2382
DB	1801	ATCTCCACCTTCTTGAAGATGGGTATTTCTTCAAGAGCACCCCATCCCTCTGAGCTACAG	1860
QY	2383	TTCAATTCACCTCTGTTTCCAGAGTCTTTTTCAGCAATGTCTATGTCTTGGAGGATGAG	2442
DB	1861	TTCAATTCACCTCTGTTTCCAGAGTCTTTTTCAGCAATGTCTATGTCTTGGAGGATGAG	1920
QY	2443	AGGGGAGGAGTAAACATTTCTAATTTGCAATAGATTTTGGAAAGACGCTAGAGCATAT	2502
DB	1921	AGGGGAGGAGTAAACATTTCTAATTTGCAATAGATTTTGGAAAGACGCTAGAGCATAT	1980
QY	2503	GGAAATACATGGGCTGTTTGGGGCATCAACACACGTTTCTTCTTGGGCTGTTAAGTGAT	2562

Db	1981	GGAAATACATGGCGCTGTTGGGCGATCAACACACAGTTTCTTATTTGGCGCTGTTAAAGTGAT	2040
QY	2563	GAGGGGAGAGAGAGATGAGAGAACTCTTTTACTGCGCGTGTCTCAGGGGAGGAACCTG	2622
Db	2041	GAGGGGAGAGAGAGATGAGAGAACTCTTTTCACTGCGCGTGTCTCAGGGGAGGAACCTG	2100
QY	2623	ATGCACTGGGTCCTGCTGAGCTGCTGTCAGCCACACTCTCTGAGTCCCTCCAC	2682
Db	2101	ATGCACTGGGTCCTGCTGAGCTGCTGTCAGCCACACTCTCTGAGTCCCTCCAC	2160
QY	2683	TGCTTTGACGAGCTCGGAACAAACGTTCTTGACAAAGTGAAGGCCCATTTGGAAGAA	2742
Db	2161	TGCTTTGACGAGCTCGGAACAAACGTTCTTGACAAAGTGAAGGCCCATTTGGAAGAA	2220
QY	2743	ATGGGCACTGTGTAGAGAAACAGACATGGAGCTCTTAGTGTGACACTTCTGCAATTAATTC	2802
Db	2221	ATGGGCACTGTGTAGAGAAACAGACATGGAGCTCTTAGTGTGACACTTCTGCAATTAATTC	2280
QY	2803	AGCGCCACGTGAAGAGCTTCAGCTGATTGAGGCGAGCAGACACAGATCAACATGGAGC	2862
Db	2281	AGCGCCACGTGAAGAGCTTCAGCTGATTGAGGCGAGCAGACACAGATCAACATGGAGC	2340
QY	2863	CCACCATGGTAGTCTCTGTTTCAAGTGGTCCAGTCAAGATGCCCTATTGGCAGATTCTC	2922
Db	2341	CCACCATGGTAGTCTCTGTTTCAAGTGGTCCAGTCAAGATGCCCTATTGGCAGATTCTC	2400
QY	2923	TTCTCCGTCTCAAGGTCAACAGAAACCTGAAAGGAGCTGGACTTAAGTGGAACTCGCTG	2982
Db	2401	TTCTCCGTCTCAAGGTCAACAGAAACCTGAAAGGAGCTGGACTTAAGTGGAACTCGCTG	2460
QY	2983	AGCCACTCTGCAAGTGAAGAGTCTTTGTAGAGCCCTGAGACGCCCTCGCTCCCTGGAG	3042
Db	2461	AGCCACTCTGCAAGTGAAGAGTCTTTGTAGAGCCCTGAGACGCCCTCGCTCCCTGGAG	2520
QY	3043	ACCCTCGGTTGGCTGGCTGTGGCTTCACAGCTGAGGACTGCAAGGACCTTGCTTTGGG	3102
Db	2521	ACCCTCGGTTGGCTGGCTGTGGCTTCACAGCTGAGGACTGCAAGGACCTTGCTTTGGG	2580
QY	3103	CTGAGAGCAACAGACCTTGACCGAGCTGGACCTGAGCTTCAATGTGCTCAAGATGCT	3162
Db	2581	CTGAGAGCAACAGACCTTGACCGAGCTGGACCTGAGCTTCAATGTGCTCAAGATGCT	2640
QY	3163	GGAGCAAAACACTTTGCCAGAGACTGAGACGCGAGCTGCAAGCTACAGGACTGCGAG	3222
Db	2641	GGAGCAAAACACTTTGCCAGAGACTGAGACGCGAGCTGCAAGCTACAGGACTGCGAG	2700
QY	3223	CTGGTCAGCTGTGGCTTCAGCTCTGACTGTGCCAGGACTGGCTCTGTGCTTAGTGCC	3282
Db	2701	CTGGTCAGCTGTGGCTTCAGCTCTGACTGTGCCAGGACTGGCTCTGTGCTTAGTGCC	2760
QY	3283	AGCCCCAGGCTGAAGAGCTAGACTGTGACGAGAAACAACTGGATGAAGCTTGGCGTGGGA	3342
Db	2761	AGCCCCAGGCTGAAGAGCTAGACTGTGACGAGAAACAACTGGATGAAGCTTGGCGTGGGA	2820
QY	3343	CTGCTCTGTGAGGGGCTAGGATCTGCTGCTGCAAACTCATACGCTGGGGCTGGACCGAG	3402
Db	2821	CTGCTCTGTGAGGGGCTAGGATCTGCTGCTGCAAACTCATACGCTGGGGCTGGACCGAG	2880
QY	3403	ACAACTCTGAGTGAAGATGAGAGAGGAACTGAGGGCCCTGAGGAGGAGAAACCTCAG	3462
Db	2881	ACAACTCTGAGTGAAGATGAGAGAGGAACTGAGGGCCCTGAGGAGGAGAAACCTCAG	2940
QY	3463	CTGCTCATCTTCAGCAGACGGAACCAAGTGTGATGACCCCTACTGAGGGCTGGATAGC	3522
Db	2941	CTGCTCATCTTCAGCAGACGGAACCAAGTGTGATGACCCCTACTGAGGGCTGGATAGC	3000
QY	3523	GGAGAGATGAGTAATAGCATCTCTCACTCAAGCGCAGAGCTCGGATCAGAGGGGCG	3582
Db	3001	GGAGAGATGAGTAATAGCATCTCTCACTCAAGCGCAGAGCTCGGATCAGAGGGGCG	3060
QY	3583	GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCTCGGAGCTGAGCAAGATCTTCCCAATT	3642

Db	3061	GCTTCCCATGTTGCTCAGGCTAACTCAAACCTCTGGACGTGAGCAAGATCTTCCCAATT	3120
QY	3643	GCTGAGATTGAGAGGAAGAGCTCCCCAGAGGTAGTACCGGTGGAACCTCTGTTGCGTGCT	3702
Db	3121	GCTGAGATTGAGAGGAAGAGCTCCCCAGAGGTAGTACCGGTGGAACCTCTGTTGCGTGCT	3180
QY	3703	TCTCTGCTCTCTCAAGGGGACCTGCATACGAAGCTTTGGGGAGCTGACGATGACTTCTGG	3762
Db	3181	TCTCTGCTCTCTCAAGGGGACCTGCATACGAAGCTTTGGGGAGCTGACGATGACTTCTGG	3240
QY	3763	GGCCCCACGGGGCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTT	3822
Db	3241	GGCCCCACGGGGCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTT	3300
QY	3823	CACTTCCCTGTAGCTGGCTCTACCGCTGGGCCAACACGGGTCTCTGCTTTGTGATGAGA	3882
Db	3301	CACTTCCCTGTAGCTGGCTCTACCGCTGGGCCAACACGGGTCTCTGCTTTGTGATGAGA	3360
QY	3883	GAAAGCGTGACCGTTGAGATTGAATTTCTGTGTGTGGGACCAAGTTCTTGGGTGAGATCAAC	3942
Db	3361	GAAAGCGTGACCGTTGAGATTGAATTTCTGTGTGTGGGACCAAGTTCTTGGGTGAGATCAAC	3420
QY	3943	CCACAGCAGCTGATGGTGGCAGGGCTCTGTCTGGAACATCAAGGCTGAGCCTGGAGCT	4002
Db	3421	CCACAGCAGCTGATGGTGGCAGGGCTCTGTCTGGAACATCAAGGCTGAGCCTGGAGCT	3480
QY	4003	GTGGAAGCTGTGCACCTCCCTCACTTTGTGTGCTCTCAAAGGGGCCATGTGGACACATCC	4062
Db	3481	GTGGAAGCTGTGCACCTCCCTCACTTTGTGTGCTCTCAAAGGGGCCATGTGGACACATCC	3540
QY	4063	CTGTTTCCAAATGGGCCACTTTTAAAGAGGAGGGAGTGTCTCTGGAGAAAGCCAGGGGTG	4122
Db	3541	CTGTTTCCAAATGGGCCACTTTTAAAGAGGAGGGAGTGTCTCTGGAGAAAGCCAGGGGTG	3600
QY	4123	GAGCTGCATCATATAGTTCTTGGAAAAACCCAGCTTCTCCCTCTGGGAGTCTCTCTGAAA	4182
Db	3601	GAGCTGCATCATATAGTTCTTGGAAAAACCCAGCTTCTCCCTCTGGGAGTCTCTCTGAAA	3660
QY	4183	ATGATCCATATGCGCTGGCTTCAATTCCTGCTCACTCTGCTGCTGCTTACCAACGC	4242
Db	3661	ATGATCCATATGCGCTGGCTTCAATTCCTGCTCACTCTGCTGCTTACCAACGC	3720
QY	4243	GTCCATCTCTGAGGAAGTCACTTCCACCTCTACCTGATCCCAAGTGAAGTCTTCCATTCGG	4302
Db	3721	GTCCATCTCTGAGGAAGTCACTTCCACCTCTACCTGATCCCAAGTGAAGTCTTCCATTCGG	3780
QY	4303	-----	4302
Db	3781	AAGGCCATAGATGATCTAGAAATGAAATTCAGATTGTCGAAATCCACAGGCCACCCCG	3840
QY	4303	-----	4302
Db	3841	CTGACCCCACTTTATATGCGGCTGTGTTTACATGTTGTTGCTGGGTCTGTTTCAAGGATGCTG	3900
QY	4303	-----AAGGAACTGGAGCTCTGCTATCGAAGCCCTGGAGAAAGACAGCTGTTT	4350
Db	3901	GAAATACCTCCCAAGGAACCTGAGCTCTGCTATCGAAGCCCTGGAGAAAGACAGCTGTTT	3960
QY	4351	TCGGAGTTCTACGTTGGCCACTTTGGGATCAGGGATCAGGCTGCAAGTGAAGAAAGAA	4410
Db	3961	TCGGAGTTCTACGTTGGCCACTTTGGGATCAGGGATCAGGCTGCAAGTGAAGAAAGAA	4020
QY	4411	GATGAGACTCTGGTGTGGAGGCTTGGTGAACACAGAGATCTCATGCTGCAACTACT	4470
Db	4021	GATGAGACTCTGGTGTGGAGGCTTGGTGAACACAGAGATCTCATGCTGCAACTACT	4080
QY	4471	CTGATCCCTCAGCCCGCATAGCCGTACTTCACTCTGGATGCCCGCAGTTGCTGAC	4530
Db	4081	CTGATCCCTCAGCCCGCATAGCCGTACTTCACTCTGGATGCCCGCAGTTGCTGAC	4140
QY	4531	TTTGTGGAACCATGATTCGAGAGCAGCTGATAGCCCGAGTGAATCGGTGGAGTTGCTTGG	4590
Db	4141	TTTGTGGAACCATGATTCGAGAGCAGCTGATAGCCCGAGTGAATCGGTGGAGTTGCTTGG	4200

QY	4591	GACAACTGCATGAGCAGGTGCTGAGCCAGGAGCAGTACGAGAGGCTGCTGGCTGAGAAC	4650
DB	4201	GACAACTGCATGAGCAGGTGCTGAGCCAGGAGCAGTACGAGAGGCTGCTGGCTGAGAAC	4260
QY	4651	ACAGGCCAGCAGCAGTGGGAAGCTGTTTCAGTTGAGCCAGTCTCGGACCGGAAGTCC	4710
DB	4261	ACAGGCCAGCAGCAGTGGGAAGCTGTTTCAGTTGAGCCAGTCTCGGACCGGAAGTCC	4320
QY	4711	AAAGATGGACTCTACCAAGCCCTGAAGGAGACCCATCTTCACTCATTATGGAACCTCTGG	4770
DB	4321	AAAGATGGACTCTACCAAGCCCTGAAGGAGACCCATCTTCACTCATTATGGAACCTCTGG	4380
QY	4771	GAGAGGGCAGCAAAAAGGACTCTCTGCCACTCAGCAGCTGA	4812
DB	4381	GAGAGGGCAGCAAAAAGGACTCTCTGCCACTCAGCAGCTGA	4422
RESULT 7			
AA02761			
ID	AA02761	standard; cDNA; 4194 BP.	
XX			
AC	AA02761;		
XX			
DT	31-MAY-2001	(first entry)	
XX			
DE	Human NB-ARC and CARD containing protein (NAC) gamma isoform cDNA.		
XX			
KW	Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;		
KW	caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;		
KW	cysteine aspartyl protease; apoptosis; cytokine production;		
KW	cytokine receptor signalling; therapy; inflammatory disorder; sepsis;		
KW	fibrosis; arthritis; cancer; adenocarcinoma; leukaemia; ss.		
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..4194	
FT		/tag= a	
FT		/product= "Human NB-ARC and CARD containing protein (NAC)	
FT		gamma isoform"	
FT	misc_feature	1..2868	
FT		/tag= b	
FT		/note= "Corresponds to 1-2868 residues of human NAC	
FT		beta isoform (AA02760)"	
FT		2869..3687	
FT	misc_feature		
FT		/tag= c	
FT		/note= "Corresponds to 2962-3780 residues of human NAC	
FT		beta isoform (AA02760)"	
FT		3688..4194	
FT	misc_feature		
FT		/tag= d	
FT		/note= "Corresponds to 3916-4422 residues of human NAC	
FT		beta isoform (AA02760)"	
XX			
PN	W0200116170-A2.		
XX			
PN	08-MAR-2001.		
XX			
PF	01-SEP-2000; 2000NO-US24152.		
XX			
PR	01-SEP-1999; 99US-0388221.		
XX			
PA	(BURN-) BURNHAM INST.		
XX			
PI	Reed JC;		
XX			
DR	WPI; 2001-183258/18.		
DR	P-PSDB; AAY72670.		
XX			
PT	Novel nucleic acid encoding NB-ARC and caspase associated recruitment		
PT	domains, used to produce polypeptides for screening for modulators of		
PT	apoptosis -		
XX			

PS	Claim 4; Page 138-143; 184pp; English.		
XX	The present sequence is a human NB-ARC and CARD containing protein		
CC	(NAC) gamma isoform cDNA. NAC gamma isoform represents the NAC splice		
CC	variant in which both the splice regions are absent in the translated		
CC	polypeptide. NAC protein comprises a nucleotide binding (NB) domain		
CC	(also referred as NB-ARC domain), a caspase-associated recruitment		
CC	domain (CARD) and a TIM-Barrel-like domain. The caspases, cysteine		
CC	aspartyl proteases, are principal effectors of apoptosis. CARD containing		
CC	NAC proteins are used for screening modulators that modulates apoptosis,		
CC	cytokine production, cytokine receptor signalling and other cellular		
CC	processes. NAC can act as an immunogen for the production of polyclonal		
CC	and monoclonal antibodies. It can also be used to diagnose and treat		
CC	inflammatory disorders such as sepsis, fibrosis and arthritis and cancer		
CC	pathologies such as adenocarcinomas and leukaemias.		
XX			
SQ	Sequence 4194 BP; 984 A; 1167 C; 1176 G; 867 T; 0 other;		
	Query Match 74.9%; Score 4078; DB 22; Length 4194;		
	Best Local Similarity 97.8%; Pred. No. 0;		
	Matches 4194; Conservative 0; Mismatches 0; Indels 96; Gaps 2;		
QY	523	ATGGCTGGCGAGCCTGGGGCGCTGGCTTACTTGGAGTCTCTGAGGAGGAGGAG 582	
DB	1	ATGGCTGGCGAGCCTGGGGCGCTGGCTTACTTGGAGTCTCTGAGGAGGAGGAG 60	
QY	583	CTGAAGGAGTCTCAGCTTCTGCTCGCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG 642	
DB	61	CTGAAGGAGTCTCAGCTTCTGCTCGCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG 120	
QY	643	ACACCGCTTCAGCCAGAGAACGATGCGATGGAGTGGCTCTGCTGCTGGTCTGAG 702	
DB	121	ACACCGCTTCAGCCAGAGAACGATGCGATGGAGTGGCTCTGCTGCTGGTCTGAG 180	
QY	703	TATGGGAGCAGGGGCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGCTGAG 762	
DB	181	TATGGGAGCAGGGGCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGCTGAG 240	
QY	763	TCACTGTGGCGCCAAAGCCAGGAGGGGAGGCGCACTCTCCCTCATTCCTCCACAGCCCA 822	
DB	241	TCACTGTGGCGCCAAAGCCAGGAGGGGAGGCGCACTCTCCCTCATTCCTCCACAGCCCA 300	
QY	823	AGTGAACCCCACTGGGCTCTCCAGCCCAACCCCTCCACCGAGTGTAAATGCCCTGG 882	
DB	301	AGTGAACCCCACTGGGCTCTCCAGCCCAACCCCTCCACCGAGTGTAAATGCCCTGG 360	
QY	883	ATCCATGAATTCGGCGGGGTGCAACCGAGGCTCAGAGAGAGGGTTTGGACAGCTG 942	
DB	361	ATCCATGAATTCGGCGGGGTGCAACCGAGGCTCAGAGAGAGGGTTTGGACAGCTG 420	
QY	943	CTGACACATCTGGACCGCTGGAGAGAAATCTCTGCTCTACTCTCTACCAAGCTCTT 1002	
DB	421	CTGACACATCTGGACCGCTGGAGAGAAATCTCTGCTCTACTCTCTCTACCAAGCTCTT 480	
QY	1003	CCAAGCTCCCGAGACCATGAGTCTCAAGCCAGGAGTCAACCAACGCCCCCAATCCACA 1062	
DB	481	CCAAGCTCCCGAGACCATGAGTCTCAAGCCAGGAGTCAACCAACGCCCCCAATCCACA 540	
QY	1063	GCAGTCTGGGAGCTGGGGATCCCACTCAGCCAGGCTAGCACCCAGAGAGAGGAG 1122	
DB	541	GCAGTCTGGGAGCTGGGGATCCCACTCAGCCAGGCTAGCACCCAGAGAGAGGAG 600	
QY	1123	GCTCTGGGACCCCAATGGCTCTGGATGAACGTCAGGAATTTACTTACACAGAAATCAGA 1182	
DB	601	GCTCTGGGACCCCAATGGCTCTGGATGAACGTCAGGAATTTACTTACACAGAAATCAGA 660	
QY	1183	GAAAGAGAGAGAGAAATCAGAGAAAGGAGGCGCCCAATGGGCGAGCGGTGGTGAACG 1242	
DB	661	GAAAGAGAGAGAGAAATCAGAGAAAGGAGGCGCCCAATGGGCGAGCGGTGGTGAACG 720	
QY	1243	CCCCCAGAGGCGCACACGAGCTTACAGCCCAACCACTAGGAGGAGGCTTCTGTGAGA 1302	
DB	721	CCCCCAGAGGCGCACACGAGCTTACAGCCCAACCACTAGGAGGAGGCTTCTGTGAGA 780	

QY 1303 GAGAGCCTCTGTTCACATGGCCCTCGAAAAATGAGGATTTTAAACCAAAATTTACACAG 1362
DB 781 GAGAGCCTCTGTTCACATGGCCCTCGAAAAATGAGGATTTTAAACCAAAATTTACACAG 840
QY 1363 CTGCTACTTTCTACAAAGACCTCACCCAGAAGCCAAAGATCCCTGGTCAAGAGAAAGCTGG 1422
DB 841 CTGCTACTTTCTACAAAGACCTCACCCAGAAGCCAAAGATCCCTGGTCAAGAGAAAGCTGG 900
QY 1423 CTTGATTTATGTGGAGAGAAATCGAGACATTTAAATTGAGATCAGAGACTTTATTTGGCCCCA 1482
DB 901 CTTGATTTATGTGGAGAGAAATCGAGACATTTAAATTGAGATCAGAGACTTTATTTGGCCCCA 960
QY 1483 GGCCTGGATACCAAGAACCTCGCATAGTCACTATCGAGGGGCTGCTGGAATTTGGGAAG 1542
DB 961 GGCCTGGATACCAAGAACCTCGCATAGTCACTATCGAGGGGCTGCTGGAATTTGGGAAG 1020
QY 1543 TCAACACTGGCCAGGAGGCTGAAGGAAGCTTGGGGAGAGGCCAGCTGTATGGGACCGC 1602
DB 1021 TCAACACTGGCCAGGAGGCTGAAGGAAGCTTGGGGAGAGGCCAGCTGTATGGGACCGC 1080
QY 1603 TTCACGATGTCTTCTACTTTAGCTGCAGAGAGCTGGCCCAAGTCCAAGTGTGTAGTCTC 1662
DB 1081 TTCACGATGTCTTCTACTTTAGCTGCAGAGAGCTGGCCCAAGTCCAAGTGTGTAGTCTC 1140
QY 1663 GCTGAGCTCATCGGAAGATGGACAGCCACTCCGGCTCCCATTTAGACAGATCCTGTCT 1722
DB 1141 GCTGAGCTCATCGGAAGATGGACAGCCACTCCGGCTCCCATTTAGACAGATCCTGTCT 1200
QY 1723 AGCCACAGAGCGGTGCTCTTCTATCTCGATGTGTAGATGAGCCAGATGGGTCTTGCAG 1782
DB 1201 AGCCACAGAGCGGTGCTCTTCTATCTCGATGTGTAGATGAGCCAGATGGGTCTTGCAG 1260
QY 1783 GAGCCGAGTTCTGAGTCTGTCTGTGCACTGAGGACAGCCACAGCCGCGGATGCACTGCTG 1842
DB 1261 GAGCCGAGTTCTGAGTCTGTCTGTGCACTGAGGACAGCCACAGCCGCGGATGCACTGCTG 1320
QY 1843 GGCAGTTTCTGGGAAACCTATCTTCCGAGGACCTCTCTGATCAGCTGGAC 1902
DB 1321 GGCAGTTTCTGGGAAACCTATCTTCCGAGGACCTCTCTGATCAGCTGGAC 1380
QY 1903 ACAGTCTGCAAGACCTCTATCTTCTTTGGAGCAGCCAGTGTGGGTAGAGTCTCTGGG 1962
DB 1381 ACAGTCTGCAAGACCTCTATCTTCTTTGGAGCAGCCAGTGTGGGTAGAGTCTCTGGG 1440
QY 1963 TTCTCTGAGTCAGCAGGAGGAATTTCTACAGATATTTACAGATGAAGGCAAGCA 2022
DB 1441 TTCTCTGAGTCAGCAGGAGGAATTTCTACAGATATTTACAGATGAAGGCAAGCA 1500
QY 2023 ATTAGAGCCTTTAGGTTGGTCAAAATCAAAACAAAGAGCTCTGGGCCCTGTCTTGGCCC 2082
DB 1501 ATTAGAGCCTTTAGGTTGGTCAAAATCAAAACAAAGAGCTCTGGGCCCTGTCTTGGCCC 1560
QY 2083 TGGGTGTCTGGCTGGCCCTGCACTTGCCTGATGACAGAGATGAAGCGGAAGAAAACTC 2142
DB 1561 TGGGTGTCTGGCTGGCCCTGCACTTGCCTGATGACAGAGATGAAGCGGAAGAAAACTC 1620
QY 2143 ACACTGACTTTCAAAGACCAACCAACCCCTCTGTCTACATTTACCTTGGCCAGGCTCTCCAA 2202
DB 1621 ACACTGACTTTCAAAGACCAACCAACCCCTCTGTCTACATTTACCTTGGCCAGGCTCTCCAA 1680
QY 2203 GCTCAGCAATGGGACCCGAGCTCAGAGACCTCTGCTCTCTGGCTCAGGGCATCTGG 2262
DB 1681 GCTCAGCAATGGGACCCGAGCTCAGAGACCTCTGCTCTCTGGCTCAGGGCATCTGG 1740
QY 2263 CAAAAAAGACCCCTTTTCAAGTCAGATGACCTCAGGAAGCATGGGTAGATGGGGCCATC 2322
DB 1741 CAAAAAAGACCCCTTTTCAAGTCAGATGACCTCAGGAAGCATGGGTAGATGGGGCCATC 1800
QY 2323 ATCTCCACTTCTTGAAGATGGGTATTTCTCAAGAGCA CCCCATCCCTCTGAGCTACAGC 2382
DB 1801 ATCTCCACTTCTTGAAGATGGGTATTTCTTCAAGAGCACCCCATCCCTCTGAGCTACAGC 1860

QY 2383 TTCAATCACCTCTGTTTCCAAAGAGTCTTTTGAGCAATGTCTATGTCTTGAGGATGAG 2442
DB 1861 TTCAATCACCTCTGTTTCCAAAGAGTCTTTTGAGCAATGTCTATGTCTTGAGGATGAG 1920
QY 2443 AAGGGAGAGGTTAAACATTTCTAAATTGATCATAGATTTTGGAAAAAGACGCTAGAAGCATAT 2502
DB 1921 AAGGGAGAGGTTAAACATTTCTAAATTGATCATAGATTTTGGAAAAAGACGCTAGAAGCATAT 1980
QY 2503 GGAATACATAGGCTGTTTGGGGCATCAACCAACAGGTTTCTATTGGGCTGTGTTAAGTGAT 2562
DB 1981 GGAATACATAGGCTGTTTGGGGCATCAACCAACAGGTTTCTATTGGGCTGTGTTAAGTGAT 2040
QY 2563 GAGGGGAGAGAGATGGAAGAACATCTTTCACTCCCGCTGTCTCAGGGAGGAAACCTG 2622
DB 2041 GAGGGGAGAGAGATGGAAGAACATCTTTCACTCCCGCTGTCTCAGGGAGGAAACCTG 2100
QY 2623 ATGCAAGTGGTCCCTGCTGCTGAGTGTGTGAGCCACACACTCTCTGAGTCCCTCCAC 2682
DB 2101 ATGCAAGTGGTCCCTGCTGCTGAGTGTGTGAGCCACACACTCTCTGAGTCCCTCCAC 2160
QY 2683 TGCTTTGACGAGCTCGGAACAAACAGTTCCTGACACAAAGTGATGGCCCATTTCCGAAGAA 2742
DB 2161 TGCTTTGACGAGCTCGGAACAAACAGTTCCTGACACAAAGTGATGGCCCATTTCCGAAGAA 2220
QY 2743 ATGGGCATGTGTAGAAAAACAGACATGAGAGTCTTTAGTGTGCACTTTCTGCATTAAATTC 2802
DB 2221 ATGGGCATGTGTAGAAAAACAGACATGAGAGTCTTTAGTGTGCACTTTCTGCATTAAATTC 2280
QY 2803 AGCCGCCACGTGAAGAACCTTTCAGCTGATTTGAGGCGACGGCAGCACAGATCAACATGGAGC 2862
DB 2281 AGCCGCCACGTGAAGAACCTTTCAGCTGATTTGAGGCGACGGCAGCACAGATCAACATGGAGC 2340
QY 2863 CCCACATGTTAGTCTCTGTTGAGTGGTCCCACTCAGATGCTCTATTTGGCAGATTTCTC 2922
DB 2341 CCCACATGTTAGTCTCTGTTGAGTGGTCCCACTCAGATGCTCTATTTGGCAGATTTCTC 2400
QY 2923 TTCTCCGTCTCAAGGTCAACAGAAACCTGAGGAGCTGGACCTTAAGTGGAAATTCGCTG 2982
DB 2401 TTCTCCGTCTCAAGGTCAACAGAAACCTGAGGAGCTGGACCTTAAGTGGAAATTCGCTG 2460
QY 2983 AGCCCATCTGCAAGTGAAGAGTCTTTGTAAGACCTTGAGACGCTCTGCTGCTCTGAG 3042
DB 2461 AGCCCATCTGCAAGTGAAGAGTCTTTGTAAGACCTTGAGACGCTCTGCTGCTCTGAG 2520
QY 3043 ACCCTGCGGTTGGCTGCTGCTGCTCAGCTGAGGACTGCAAGGACTCTTGCTTTGGG 3102
DB 2521 ACCCTGCGGTTGGCTGCTGCTGCTCAGCTGAGGACTGCAAGGACTCTTGCTTTGGG 2580
QY 3103 CTGAGAGCCAAACAGACCCCTGACCGAGCTGGAACCTTGAGCTTCAATGTGCTCACGGATGCT 3162
DB 2581 CTGAGAGCCAAACAGACCCCTGACCGAGCTGGAACCTTGAGCTTCAATGTGCTCACGGATGCT 2640
QY 3163 GSAGCCAAACACCTTTGCCAGAGACTGAGACGCGAGCTGCAAGCTACAGCGACTGAG 3222
DB 2641 GSAGCCAAACACCTTTGCCAGAGACTGAGACGCGAGCTGCAAGCTACAGCGACTGAG 2700
QY 3223 CTGGTCAGCTGGGCTCAGCTGCTGCTGCTGCCAGGACCTGGCCCTCTGCTGTTAGTGCC 3282
DB 2701 CTGGTCAGCTGGGCTCAGCTGCTGCTGCTGCCAGGACCTGGCCCTCTGCTGTTAGTGCC 2760
QY 3283 AGCCCCAGCTGAAGGAGCTAGACCTGCAGCAGAAACCTTGGATGAGCTTTGGCGTGGGA 3342
DB 2761 AGCCCCAGCTGAAGGAGCTAGACCTGCAGCAGAAACCTTGGATGAGCTTTGGCGTGGGA 2820
QY 3343 CTGCTCTGTGAGGCTCAGGATCCTGCTGCTGCAAACTCATACGCTCGGGGTGGAACGAG 3402
DB 2821 CTGCTCTGTGAGGCTCAGGATCCTGCTGCTGCAAACTCATACGCTCTGCTGTTAGTGCC 2867
QY 3403 ACACCTCTGATGATGAGATGAGGAGGAACTGAGGGCCCTGGAGCAGGAGAAACCTCAG 3462
DB 2868 ----- 2867
QY 3463 CTGCTCATCTTCAGCAGACGGAAACCAAGTGTGATGACCCCTTACTGAGGGCCTGGATACG 3522

Db 2868 -----GAAACCAAGTGTGATGACCCCTACTGAGGGCTGTGATACG 2907
 QY 3523 GGAGAGATGAGTAATAGACATCTCTCACTCAAGCGCGAGAGACTCGGATCAGAGAGGGG 3582
 Db 2908 GGAGAGATGAGTAATAGACATCTCTCACTCAAGCGCGAGAGACTCGGATCAGAGAGGGG 2967
 QY 3583 GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCTCGAGCTGAGCAAGATCTTCCCAATT 3642
 Db 2968 GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCTCGAGCTGAGCAAGATCTTCCCAATT 3027
 QY 3643 GCTGAGATGACAGAGAAAGCTCCCGAGAGGTAGTACCGGTGGAACCTTTGTCGCTGCT 3702
 Db 3028 GCTGAGATGACAGAGAAAGCTCCCGAGAGGTAGTACCGGTGGAACCTTTGTCGCTGCT 3087
 QY 3703 TCTCTGCTCTCAAGGGGACCTGATACCAAGCCCTTTGGGACCTGACGATGACTTCTGG 3762
 Db 3088 TCTCTGCTCTCAAGGGGACCTGATACCAAGCCCTTTGGGACCTGACGATGACTTCTGG 3147
 QY 3763 GGCCCCACGGGCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTTGACCGAGTT 3822
 Db 3148 GGCCCCACGGGCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTTGACCGAGTT 3207
 QY 3823 CACTTCCCTGTAGTGGCTCTACCGCTGGCCCAACACGGGTCTCTGCTTTGTGATGAGA 3882
 Db 3208 CACTTCCCTGTAGTGGCTCTACCGCTGGCCCAACACGGGTCTCTGCTTTGTGATGAGA 3267
 QY 3883 GAAGCGGTGACCGTTGAGATTGCAATTTCTGTGTGGGACCACTCTGGGTGAGATCAAC 3942
 Db 3268 GAAGCGGTGACCGTTGAGATTGCAATTTCTGTGTGGGACCACTCTGGGTGAGATCAAC 3327
 QY 3943 CCACAGCAGCTGGATGTGGCAGGGCTCTGCTGGACATCAAGCTGAGCTGGAGCT 4002
 Db 3328 CCACAGCAGCTGGATGTGGCAGGGCTCTGCTGGACATCAAGCTGAGCTGGAGCT 3387
 QY 4003 GTGGAAGCTGTGACCTCCCTCACTTTGTGGCTCTCCAAAGGGGCGCATGTGACACATCC 4062
 Db 3388 GTGGAAGCTGTGACCTCCCTCACTTTGTGGCTCTCCAAAGGGGCGCATGTGACACATCC 3447
 QY 4063 CTGTTCCAAATGGCCACATTTAAAGAGGAGGGATGCTCTCTGAGAGCCAGCCAGGGTG 4122
 Db 3448 CTGTTCCAAATGGCCACATTTAAAGAGGAGGGATGCTCTCTGAGAGCCAGCCAGGGTG 3507
 QY 4123 GAGCTGCATCAGATGTTCTGGAACACCCAGCTTCTCCCTCTGGGAGTCTCTCTGAAA 4182
 Db 3508 GAGCTGCATCAGATGTTCTGGAACACCCAGCTTCTCCCTCTGGGAGTCTCTCTGAAA 3567
 QY 4183 ATGATCCATAATGCTGCTGCTCAATCCCGTCACTCTGTGGTGTGTTTACCAACCGC 4242
 Db 3568 ATGATCCATAATGCTGCTGCTCAATCCCGTCACTCTGTGGTGTGTTTACCAACCGC 3627
 QY 4243 GTCCATCCTGAGGAGTCACTTCCACCTCTACCTGATCCCAAGTGAAGTCTCTCCATTGG 4302
 Db 3628 GTCCATCCTGAGGAGTCACTTCCACCTCTACCTGATCCCAAGTGAAGTCTCTCCATTGG 3686
 QY 4303 AAGGAATCGAGCTCTGCTATCAAGCCCTGGAAGAGACAGCTGTTCTCGAGTCTTAC 4362
 Db 3687 --GGAATCGAGCTCTGCTATCAAGCCCTGGAAGAGACAGCTGTTCTCGAGTCTTAC 3744
 QY 4363 GTTGGCCACTTGGGATCAGGGATCAGGCTGCAAGTGAAGAGCAAGAAAGATGAGACTCTG 4422
 Db 3745 GTTGGCCACTTGGGATCAGGGATCAGGCTGCAAGTGAAGAGCAAGAAAGATGAGACTCTG 3804
 QY 4423 GTTGGGAGGCTTGTGTAACACAGAGATCTCATGCTGCAACTACTCTGATCCCTCCA 4482
 Db 3805 GTTGGGAGGCTTGTGTAACACAGAGATCTCATGCTGCAACTACTCTGATCCCTCCA 3864
 QY 4483 GCCCGCATAGCGTACTCTTCACTCTGGATGCCCGCAGTTGCTGACTTTGTGGACAG 4542
 Db 3865 GCCCGCATAGCGTACTCTTCACTCTGGATGCCCGCAGTTGCTGACTTTGTGGACAG 3924
 QY 4543 TATCGAGACAGCTGATAGCCCGAGTGACATCGGTGAGGTTGCTTTGGACAACTGCAT 4602

Db 3925 TATCGAGAGCAGCTGATAGCCCGAGTGACATCGGTGGAGGTTGTCTTGACAAACTGCAT 3984
 QY 4603 GGACAGGTGCTGAGCCAGGAGCAGTACGAGAGGGTGTCTGGCTGAGAACACGAGGCCAGC 4662
 Db 3985 GGACAGGTGCTGAGCCAGGAGCAGTACGAGAGGGTGTCTGGCTGAGAACACGAGGCCAGC 4044
 QY 4663 CAGATCGGAGAGCTGTTGAGCTTGAGCCAGTCTCTGGACCGGAAGTGAAGATGAGACTC 4722
 Db 4045 CAGATCGGAGAGCTGTTGAGCTTGAGCCAGTCTCTGGACCGGAAGTGAAGATGAGACTC 4104
 QY 4723 TACCAAGCCCTGAAGGAGACCATCTCTCACTCATTTATGGAACCTCTGGAGAGGGCAGC 4782
 Db 4105 TACCAAGCCCTGAAGGAGACCATCTCTCACTCATTTATGGAACCTCTGGAGAGGGCAGC 4164
 QY 4783 AAAAAGGAGACTCTGCCACTCAGCAGCTGA 4812
 Db 4165 AAAAAGGAGACTCTGCCACTCAGCAGCTGA 4194

RESULT 8
 AAD02762
 ID AAD02762 standard; cDNA; 4329 BP.
 XX
 AC AAD02762;
 XX
 DT 31-MAY-2001 (first entry)
 XX
 DE Human NB-ARC and CARD containing protein (NAC) delta isoform cDNA.
 XX
 KW Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;
 KW caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;
 KW cysteine-aspartyl protease; apoptosis; cytokine production;
 KW cytokine receptor signalling; therapy; inflammatory disorder; sepsis;
 KW fibrosis; arthritis; cancer; adenocarcinoma; leukaemia; ss.
 XX
 OS Homo sapiens.
 XX
 FT Key
 FT CDS
 FT Location/Qualifiers
 FT 1..4329
 FT /tag= a
 FT /product= "Human NB-ARC and CARD containing protein (NAC)
 FT delta isoform"
 FT 1..2868
 FT misc_feature
 FT /tag= b
 FT /note= "Corresponds to 1-2868 residues of human NAC
 FT beta isoform (AAD02760)"
 FT 2869..4329
 FT /tag= c
 FT /note= "Corresponds to 2962-4422 residues of human NAC
 FT beta isoform (AAD02760)"
 XX
 PN WO200116170-A2.
 XX
 PD 08-MAR-2001.
 XX
 PF 01-SEP-2000; 2000WO-US24152.
 XX
 PR 01-SEP-1999; 99US-0388221.
 XX
 PA (BURN-) BURNHAM INST.
 XX
 PI Reed JC;
 XX
 DR WPI; 2001-183258/18.
 DR P-PSDB; AAY72671.
 XX
 PT Novel nucleic acid encoding NB-ARC and caspase associated recruitment
 PT domains, used to produce polypeptides for screening for modulators of
 PT apoptosis -
 XX
 PS Claim 4; Page 148-154; 184pp; English.
 XX
 CC The present sequence is a human NB-ARC and CARD containing protein

CC (NAC) delta isoform cDNA. NAC delta isoform represents the NAC splice
 CC variant in which one of the splice region is absent in the translated
 CC polypeptide. NAC protein comprises a nucleotide binding (NB) domain
 CC (also referred as NB-ARC domain), a caspase-associated recruitment
 CC domain (CARD) and a TRIM-Barrel-like domain. The caspases, cysteine
 CC aspartyl proteases, are principal effectors of apoptosis. CARD containing
 CC NAC proteins are used for screening modulators that modulates apoptosis,
 CC cytokine production, cytokine receptor signalling and other cellular
 CC processes. NAC can act as an immunogen for the production of polyclonal
 CC and monoclonal antibodies. It can also be used to diagnose and treat
 CC inflammatory disorders such as sepsis, fibrosis and arthritis and cancer
 CC pathologies such as adenocarcinomas and leukaemias.
 XX
 SQ Sequence 4329 BP; 1018 A; 1202 C; 1208 G; 901 T; 0 other;

Query Match 72.6%; Score 3952; DB 22; Length 4329;
 Best Local Similarity 94.9%; Pred. No. 0;
 Matches 4197; Conservative 0; Mismatches 0; Indels 225; Gaps 2;

QY 523 ATGGCTGGCGGAGCTGGGGCGCCCTGCTTACTTGGAGTTCCTGAGAGGAGGAG 582
 DB 1 ATGGCTGGCGGAGCTGGGGCGCCCTGCTTACTTGGAGTTCCTGAGAGGAGGAG 60
 QY 583 CTGAAGGAGTTCAGCTTCTGCTCGCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG 642
 DB 61 CTGAAGGAGTTCAGCTTCTGCTCGCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG 120
 QY 643 ACACCGCTCAGCCAGAGAGAGAGTGGATGGAGTGGCTCGTACCTGGTGGCTCAG 702
 DB 121 ACACCGCTCAGCCAGAGAGAGAGTGGATGGAGTGGCTCGTACCTGGTGGCTCAG 180
 QY 703 TATGGGAGACGGGCGCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGCTGAGG 762
 DB 181 TATGGGAGACGGGCGCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGCTGAGG 240
 QY 763 TCACTGTGGCGCCCAAGCCAGAGAGGGGAGGCGCACTCTCCCTCATTCCTCCCTACAGCCCA 822
 DB 241 TCACTGTGGCGCCCAAGCCAGAGAGGGGAGGCGCACTCTCCCTCATTCCTCCCTACAGCCCA 300
 QY 823 AFTGAACCCCACTGGGGTCTCCAGCCAAACCCCTCAGCGAGTGTAATGCCCTGG 882
 DB 301 AFTGAACCCCACTGGGGTCTCCAGCCAAACCCCTCAGCGAGTGTAATGCCCTGG 360
 QY 883 ATCCATGAATTCGGCGGGGTGCACCCAGGCTCAGAGAGAGGTTTGGAGACGCTG 942
 DB 361 ATCCATGAATTCGGCGGGGTGCACCCAGGCTCAGAGAGAGGTTTGGAGACGCTG 420
 QY 943 CTTGACACATCTGGAGCGCGCTGGAGAGAAATCTCTGCTCCTCCTACCAAGCTCTT 1002
 DB 421 CTTGACACATCTGGAGCGCGCTGGAGAGAAATCTCTGCTCCTCCTACCAAGCTCTT 480
 QY 1003 CCAAGCTCCCCAGACCATGAGTCTCAAGCCAGGAGTCAACCAACGCCCCCAATCCACA 1062
 DB 481 CCAAGCTCCCCAGACCATGAGTCTCAAGCCAGGAGTCAACCAACGCCCCCAATCCACA 540
 QY 1063 GCAGTCTGGGAGCTGGGAGTCCCACTCAGCCAGCCTAGCACCCAGAGCAGGAG 1122
 DB 541 GCAGTCTGGGAGCTGGGAGTCCCACTCAGCCAGCCTAGCACCCAGAGCAGGAG 600
 QY 1123 GCTCTGGGACCCCAATGGCTCTGGATGAAAAGTCAAGAAATTTACTTACACAGAAATCAGA 1182
 DB 601 GCTCTGGGACCCCAATGGCTCTGGATGAAAAGTCAAGAAATTTACTTACACAGAAATCAGA 660
 QY 1183 GAAAGAGAGAGAGAGAAATCAGAGAAAGGAGGCGCCCAATGGGAGCGGTGTAGGAAAG 1242
 DB 661 GAAAGAGAGAGAGAGAAATCAGAGAAAGGAGGCGCCCAATGGGAGCGGTGTAGGAAAG 720
 QY 1243 CCCCCACAGCGCACACCGCTTACAGCCCCCACCACCAACCCATGGGAGCTTCTGTGAGA 1302
 DB 721 CCCCCACAGCGCACACCGCTTACAGCCCCCACCACCAACCCATGGGAGCTTCTGTGAGA 780
 QY 1303 GAGAGCTCTGTTCACATAGGCGCTCGAAAAATGAGGATTTTAAACCAAAATTCACACAG 1362

DB 781 GAGAGCTCTGTTCACATGCGCCCTGGAAAAATGAGGATTTTAAACCAAAATTCACACAG 840
 QY 1363 CTGCTACTTCTACAAAGACCTCACCCAGAGAGCCAGATCCCTGGTCAAGAGAGAGCTGG 1422
 DB 841 CTGCTACTTCTACAAAGACCTCACCCAGAGAGCCAGATCCCTGGTCAAGAGAGAGCTGG 900
 QY 1423 CTTGATTTATGTGGAGAGAGATCGAGGACATTTAATTGAGATCAGAGACTTATTTGGCCCA 1482
 DB 901 CTTGATTTATGTGGAGAGAGATCGAGGACATTTAATTGAGATCAGAGACTTATTTGGCCCA 960
 QY 1483 GGCCTGGATACCCAGAAACCTTCGATATCATATGCAAGGGGCTGCTGGAATTTGGGAAG 1542
 DB 961 GGCCTGGATACCCAGAAACCTTCGATATCATATGCAAGGGGCTGCTGGAATTTGGGAAG 1020
 QY 1543 TCNAACACTGGCCAGGCTGNAAGGAGCCTGGGGAGAGCCAGCTATGTTGGGACCGC 1602
 DB 1021 TCNAACACTGGCCAGGCTGNAAGGAGCCTGGGGAGAGCCAGCTATGTTGGGACCGC 1080
 QY 1603 TTCCAGCATGTCTTCTACTTTCAGCTTCAGAGAGCTGGCCCAAGTCCAAAGTGGTGAAGTCTC 1662
 DB 1081 TTCCAGCATGTCTTCTACTTTCAGCTTCAGAGAGCTGGCCCAAGTCCAAAGTGGTGAAGTCTC 1140
 QY 1663 GCTGAGCTCATCGGAAAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCCTGTCT 1722
 DB 1141 GCTGAGCTCATCGGAAAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCCTGTCT 1200
 QY 1723 AGGCCAGAGCGGCTCTTCTTCTATCTCGATGGTGTAGATGAGCCAGGATGGGTCTTGCGAG 1782
 DB 1201 AGGCCAGAGCGGCTCTTCTTCTATCTCGATGGTGTAGATGAGCCAGGATGGGTCTTGCGAG 1260
 QY 1783 GAGCCGAGTTCTGAGCTCTGTCTGCACTGGAGCCAGCCACAGCCGCGCGGATGCACTGCTG 1842
 DB 1261 GAGCCGAGTTCTGAGCTCTGTCTGCACTGGAGCCAGCCACAGCCGCGCGGATGCACTGCTG 1320
 QY 1843 GGCAGTTTCTGGGAAAAAATATATCTCCGAGGCACTCTTCTCTGATCAGCGCTCGGACC 1902
 DB 1321 GGCAGTTTCTGGGAAAAAATATATCTCCGAGGCACTCTTCTCTGATCAGCGCTCGGACC 1380
 QY 1903 ACAGCTCTGCAGAACCTTCTTCTTTGGAGCAGGCACTGGGTAGAGTCTCTGGGG 1962
 DB 1381 ACAGCTCTGCAGAACCTTCTTCTTTGGAGCAGGCACTGGGTAGAGTCTCTGGGG 1440
 QY 1963 TTCTCTGAGTCTCAGCAGGAAAGGAAATATTTCTACAGATATTTTACAGATGAAAGGCAAGCA 2022
 DB 1441 TTCTCTGAGTCTCAGCAGGAAAGGAAATATTTCTACAGATATTTTACAGATGAAAGGCAAGCA 1500
 QY 2023 ATTAGAGCTTTAGGTTGGTCAAAATCAAAAGAGCTCTGGGCGCTGTGTCTTGTGCC 2082
 DB 1501 ATTAGAGCTTTAGGTTGGTCAAAATCAAAAGAGCTCTGGGCGCTGTGTCTTGTGCC 1560
 QY 2083 TGGGTCTCTGGCTGGCTGCATTTGCTGATGCGAGCAGATGAAAGGAAAGAAATC 2142
 DB 1561 TGGGTCTCTGGCTGGCTGCATTTGCTGATGCGAGCAGATGAAAGGAAAGAAATC 1620
 QY 2143 ACAGCTGACTTCCAAAGCACCACCAACCCCTCTGTCTTACATTTACCTTTGCCCAGGCTCTCCAA 2202
 DB 1621 ACAGCTGACTTCCAAAGCACCACCAACCCCTCTGTCTTACATTTACCTTTGCCCAGGCTCTCCAA 1680
 QY 2203 GCTCAGCATTGGGACCCAGCTCAGACCTCTCTCTCTGGCTGTGGGTGCTGAGGGCATCTGG 2262
 DB 1681 GCTCAGCATTGGGACCCAGCTCAGACCTCTCTCTCTGGCTGTGGGTGCTGAGGGCATCTGG 1740
 QY 2263 CAAAAAAGACCCCTTTTTCAGTCCAGATGACCTCAGAGAGCATGGGTAGATGGGGCCATC 2322
 DB 1741 CAAAAAAGACCCCTTTTTCAGTCCAGATGACCTCAGAGAGCATGGGTAGATGGGGCCATC 1800
 QY 2323 ATCTCCACCTTCTTGAAGATGGGTATTTCTTCAAGAGCACCCTCATCTCTGAGCTACAGC 2382
 DB 1801 ATCTCCACCTTCTTGAAGATGGGTATTTCTTCAAGAGCACCCTCATCTCTGAGCTACAGC 1860
 QY 2383 TTTCAATTCACCTGTGTTCCAAAGAGTCTTTGAGCAAAATGTCTATGTCTTGGAGGATGAG 2442
 DB 1861 TTTCAATTCACCTGTGTTCCAAAGAGTCTTTTGGAGCAATGTCTATGTCTTGGAGGATGAG 1920

Db	3988	CTGATCCCTCCAGCCCGCATAGCCGTACCTTCACCTCTGGATGCCCGCAGTTGCTGCAC	4047
Qy	4531	TTTGTGGACCAAGTATCGAGAGCAGCTGATAGCCGAGTGACATCGGTGGAGGTTGCTTG	4590
Db	4048	TTTGTGGACCAAGTATCGAGAGCAGCTGATAGCCGAGTGACATCGGTGGAGGTTGCTTG	4107
Qy	4591	GACAACTGATCGACACAGGTGCTGAGCCAGGACGAGTACGAGAGGGTGCTGGCTGAGAAC	4650
Db	4108	GACAACTGATCGACACAGGTGCTGAGCCAGGACGAGTACGAGAGGGTGCTGGCTGAGAAC	4167
Qy	4651	ACGAGGCCAGCCAGATGCGGAAGCTTTTTCAGTTTGAGCCAGTCTCTGGGACCGGAAGTGC	4710
Db	4168	ACGAGGCCAGCCAGATGCGGAAGCTTTTTCAGTTTGAGCCAGTCTCTGGGACCGGAAGTGC	4227
Qy	4711	AAAGATGGAGCTTACCAAGCCCTGAAGGAGACCCATCCTCACCTCATATTGGAACCTCTGG	4770
Db	4228	AAAGATGGAGCTTACCAAGCCCTGAAGGAGACCCATCCTCACCTCATATTGGAACCTCTGG	4287
Qy	4771	GAGAAGGGCAGCAAAAAGGGAGTCTCTGCCACTCAGCAGCTGA	4812
Db	4288	GAGAAGGGCAGCAAAAAGGGAGTCTCTGCCACTCAGCAGCTGA	4329
RESULT 9			
AAD02764			
ID	AAD02764 standard; cDNA; 4556 BP.		
XX			
AC	AAD02764;		
XX			
AC	31-MAY-2001 (first entry)		
XX			
DE	Human NAC beta isoform-CARD-X1 chimeric cDNA.		
XX			
KW	Human; NB domain; nucleotide binding domain; NB-ARC and CARD; NAC;		
KW	caspase-associated recruitment domain; CARD; TIM-Barrel-like domain;		
KW	cysteine aspartyl protease; apoptosis; cytokine production;		
KW	cytokine receptor signalling; therapy; inflammatory disorder; sepsis;		
KW	fibrosis; arthritis; cancer; adenocarcinoma; leukaemia;		
KW	chimeric protein; ss.		
XX			
OS	Homo sapiens.		
XX			
Key	Location/Qualifiers		
FH	1..4365		
FT	/*tag= a		
FT	/product= "Human NAC beta isoform-CARD-X1 chimera"		
FT	/*misc_feature		
FT	1..3234		
FT	/*tag= b		
FT	/note= "Corresponds to 1-3234 residues of human		
FT	NAC beta isoform (AAD02760)"		
FT	/*misc_feature		
FT	3235..4362		
FT	/*tag= c		
FT	/note= "Corresponds to 166-1293 residues of human		
FT	CARD-X (AAD02763)"		
XX			
PN	WO200116170-A2.		
XX			
PD	08-MAR-2001.		
XX			
PF	01-SEP-2000; 2000WO-US24152.		
XX			
PR	01-SEP-1999; 99US-0388221.		
XX			
PA	(BURN-) BURNHAM INST.		
XX			
PI	Reed JC;		
XX			
PI	WPI; 2001-183258/18.		
DR	P-PSDB; AAY72673.		
XX			
PT	Novel nucleic acid encoding NB-ARC and caspase associated recruitment		
PT	domains, used to produce polypeptides for screening for modulators of		

Db 721 |||||CCCCACAGCGCACACAGCTACAGCCCCACACACCCATGGAGCCTTCTGTGAGA 780
Qy 1303 GAGAGCCTCTGTTCCACATGGCCCTGGAAAAATGAGATTTTAAACCAAAAATTCACACAG 1362
Db 781 GAGAGCCTCTGTTCCACATGGCCCTGGAAAAATGAGATTTTAAACCAAAAATTCACACAG 840
Qy 1363 CTGCTACTTCTTACAAAGACCTCACCCAGAACCCAGATGCCCTGGTCAAGAGAGCTGG 1422
Db 841 CTGCTACTTCTTACAAAGACCTCACCCAGAACCCAGATGCCCTGGTCAAGAGAGCTGG 900
Qy 1423 CTTGATTATGTGGAGGAGAAATCGAGACATTTAAATTGAGATCAGAGACTTATTTGGCCCCA 1482
Db 901 CTTGATTATGTGGAGGAGAAATCGAGACATTTAAATTGAGATCAGAGACTTATTTGGCCCCA 960
Qy 1483 GGCCTGGATACCAAGAACCTCGCATAGTCATCTGACGGGGCTGCTGGAAATGGGAAG 1542
Db 961 GGCCTGGATACCAAGAACCTCGCATAGTCATCTGACGGGGCTGCTGGAAATGGGAAG 1020
Qy 1543 TCAACACTGGCCAGGAGGTGAAGGAAGCTGGGGAGAGGCCAGCTGATGGGACCGC 1602
Db 1021 TCAACACTGGCCAGGAGGTGAAGGAAGCTGGGGAGAGGCCAGCTGATGGGACCGC 1080
Qy 1603 TTCCAGCATGCTTCTTACTTACGTCGAGAGAGCTGGCCCCAGTCCAAAGGTGGTGAAGTCTC 1662
Db 1081 TTCCAGCATGCTTCTTACTTACGTCGAGAGAGCTGGCCCCAGTCCAAAGGTGGTGAAGTCTC 1140
Qy 1663 GCTGAGCTCATCGGAAAGATGGGACAGCCACTCGGGCTCCCATTAGACAGATCCTGTCT 1722
Db 1141 GCTGAGCTCATCGGAAAGATGGGACAGCCACTCGGGCTCCCATTAGACAGATCCTGTCT 1200
Qy 1723 AGGCCAGAGCGGCTGCTCTTCACTCTCGATGGTGTAGATGAGCCAGGATGGGTCTTTGCGAG 1782
Db 1201 AGGCCAGAGCGGCTGCTCTTCACTCTCGATGGTGTAGATGAGCCAGGATGGGTCTTTGCGAG 1260
Qy 1783 GAGCCGAGTTCTGAGCTCTGTCTGCACTGGAGCAGCCACAGCCGCGGATGCACTGTCTG 1842
Db 1261 GAGCCGAGTTCTGAGCTCTGTCTGCACTGGAGCAGCCACAGCCGCGGATGCACTGTCTG 1320
Qy 1843 GGCAGTTTCTGGGAAACATACTTCCGAGGCATCTTCTGATCAGCCTGGAC 1902
Db 1321 GGCAGTTTCTGGGAAACATACTTCCGAGGCATCTTCTGATCAGCCTGGAC 1380
Qy 1903 ACAGCTCTGCAGAACCTCATCTTCTTTGGAGCAGGCACGTTGGGTAGAGGTCTCTGGGG 1962
Db 1381 ACAGCTCTGCAGAACCTCATCTTCTTTGGAGCAGGCACGTTGGGTAGAGGTCTCTGGGG 1440
Qy 1963 TTCTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTACAGATGAAAGGCAAGCA 2022
Db 1441 TTCTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTACAGATGAAAGGCAAGCA 1500
Qy 2023 ATTAGAGCCTTTAGGTTGGTTCAAATCAACAAAGAGCTCTGGGCCCTGTCTGTGGCCC 2082
Db 1501 ATTAGAGCCTTTAGGTTGGTTCAAATCAACAAAGAGCTCTGGGCCCTGTGTCTGTGGCCC 1560
Qy 2083 TGGGTCTCTGGCTGGCCCTGCACTTGCCTGATCAGCAGATGAAGGGGAAGGAAAACTC 2142
Db 1561 TGGGTCTCTGGCTGGCCCTGCACTTGCCTGATGTCAGCAGATGAGCGGAGGAAAACTC 1620
Qy 2143 ACATGACTTCCAGACCAACCAACCTCTGTCTTACATTACCTTGGCCAGGCTCTCCAA 2202
Db 1621 ACATGACTTCCAGACCAACCAACCTCTGTCTTACATTACCTTGGCCAGGCTCTCCAA 1680
Qy 2203 GCTCAGCATTTGGGACCCAGCTCAGAGACCTCTGCTCTGTGGCTGCTGAGGGCATCTGG 2262
Db 1681 GCTCAGCATTTGGGACCCAGCTCAGAGACCTCTGCTCTGTGGCTGCTGAGGGCATCTGG 1740
Qy 2263 CAAAAAAGACCTTTTTCAGTCCAGATGACCTCAGGAGCATGGGTTAGATGGGGCATC 2322
Db 1741 CAAAAAAGACCTTTTTCAGTCCAGATGACCTCAGGAGCATGGGTTAGATGGGGCATC 1800
Qy 2323 ATCTCCACCTTCTTGAAGATGGGTATTTCTCAAGAGCACCCCATCTCTGAGCTACAGC 2382

Db 1801 ATCTCCACCTTCTTGAAGATGGGTATTTCTTCAAGAGCACCCCATCTCTGAGCTACAGC 1860
Qy 2383 TTCAATTCACTCTGTCTTCCAAAGAGTTCTTTTCAGCAATGTCTCTATGTCTTTGAGGATGAG 2442
Db 1861 TTCAATTCACTCTGTCTTCCAAAGAGTTCTTTTCAGCAATGTCTCTATGTCTTTGAGGATGAG 1920
Qy 2443 AAGGGGAGAGGTAAACATTTCTAAATTGCAATCATAGATTTTGGAAAAAGACGCTAGAAGCATAT 2502
Db 1921 AAGGGGAGAGGTAAACATTTCTAAATTGCAATCATAGATTTTGGAAAAAGACGCTAGAAGCATAT 1980
Qy 2503 GGAATACATGCGCTGTTTGGGGCATCAACCAACAGTTTCTTATTGGGCTGTGTTAAGTGAT 2562
Db 1981 GGAATACATGCGCTGTTTGGGGCATCAACCAACAGTTTCTTATTGGGCTGTGTTAAGTGAT 2040
Qy 2563 GAGGGGAGAGAGATGGAGAACATCTTTCACTGCGGGCTGTCTCAGGGGAGGAACTCG 2622
Db 2041 GAGGGGAGAGAGATGGAGAACATCTTTCACTGCGGGCTGTCTCAGGGGAGGAACTCG 2100
Qy 2623 ATGCAAGTGGTCCCGTCCCTGAGCTGTCTGAGCCACACTCTCTGAGTGCCTCCAC 2682
Db 2101 ATGCAAGTGGTCCCGTCCCTGAGCTGTCTGAGCCACACTCTCTGAGTGCCTCCAC 2160
Qy 2683 TGCTTGTACGAGACTCGGAACAAACAGTTCTTGACACAAAGTGATGGCCCATTTCCGAAGAA 2742
Db 2161 TGCTTGTACGAGACTCGGAACAAACAGTTCTTGACACAAAGTGATGGCCCATTTCCGAAGAA 2220
Qy 2743 ATGGGCATGTGTAGAAAAAGACATGGAGCTCTTAGTGTGCATTTTCTGCAATTAATTC 2802
Db 2221 ATGGGCATGTGTAGAAAAAGACATGGAGCTCTTAGTGTGCATTTTCTGCAATTAATTC 2280
Qy 2803 AGCCGCCACGTGAAGAACCTTCACTGATTTGAGGCGAGGACAGACAGATCAACATGGAGC 2862
Db 2281 AGCCGCCACGTGAAGAACCTTCACTGATTTGAGGCGAGGACAGACAGATCAACATGGAGC 2340
Qy 2863 CCCACCATGGTATGCTCTTCAAGTGGTCCCGACAGATGCCATTTGGCAGATTTCTC 2922
Db 2341 CCCACCATGGTATGCTCTTCAAGTGGTCCCGACAGATGCCATTTGGCAGATTTCTC 2400
Qy 2923 TTCTCCGCTCTCAAGGTCACAGAAACCTGAAAGAGCTGGACCTAAGTGGAAACCTCGCTG 2982
Db 2401 TTCTCCGCTCTCAAGTCAACAGAAACCTGAAAGAGCTGGACCTAAGTGGAAACCTCGCTG 2460
Qy 2983 AGCCACTCTGCAAGTGAAGAGTCTTTGTAAGACCTTGAGACGCCCTCGCTGCTCTCGGAG 3042
Db 2461 AGCCACTCTGCAAGTGAAGAGTCTTTGTAAGACCTTGAGACGCCCTCGCTGCTCTCGGAG 2520
Qy 3043 ACCCTCGGTTGGCTGGCTGTGGCCCTCAGCTGAGGACTGCAAGGACCTTGCCCTTTGGG 3102
Db 2521 ACCCTCGGTTGGCTGGCTGTGGCCCTCAGCTGAGGACTGCAAGGACCTTGCCCTTTGGG 2580
Qy 3103 CTGAGAGCCAAACAGACCTTGACCGAGCTGGACCTGAGCTTCAATGTGTCTCACGGATGCT 3162
Db 2581 CTGAGAGCCAAACAGACCTTGACCGAGCTGGACCTGAGCTTCAATGTGTCTCACGGATGCT 2640
Qy 3163 GGAGCCAAAACACTTTTGGCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCGACTGCGAG 3222
Db 2641 GGAGCCAAAACACTTTTGGCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCGACTGCGAG 2700
Qy 3223 CTGGTCAAGTGTGGCTCAGCTGTGCTGCTGCGAGGACCTGGCCCTCTGTGCTTAGTGCC 3282
Db 2701 CTGGTCAAGTGTGGCTCAGCTGTGCTGCTGCGAGGACCTGGCCCTCTGTGCTTAGTGCC 2760
Qy 3283 AGCCCCAGCCTCAAGGAGCTAGACCTGCGAGCAGAAACAACTCGATGAGCTTGGCGTGCGA 3342
Db 2761 AGCCCCAGCCTCAAGGAGCTAGACCTGCGAGCAGAAACAACTCGATGAGCTTGGCGTGCGA 2820
Qy 3343 CTGCTCTGTGAGGGCTCAGGCATCTGCTCTGCAAACTCATACGCTGTGGGCTTGGACCCAG 3402
Db 2821 CTGCTCTGTGAGGGCTCAGGCATCTGCTCTGCAAACTCATACGCTGTGGGCTTGGACCCAG 2880
Qy 3403 ACAACTCTGAGTGTAGATGAGGAGGAACTGAGGGGCCCTGGAGCAGAGAAACCTCAG 3462
Db 2881 ACAACTCTGAGTGTAGATGAGGAGGAACTGAGGGGCCCTGGAGCAGAGAAACCTCAG 2940

QY	703	TATGGGGAGCAGCGGGCTTGGACCTAGCCCTCCATATCCTGGGAGCAGATGGGGCTGAGG	762	QY	1783	GAGCCGAGTTCTGAGCTCTGTGCACTGGAGCCAGCCACAGCCGGCGGATGCATCTGCTG	1842
DB	181	TATGGGGAGCAGCGGGCTTGGACCTAGCCCTCCATATCCTGGGAGCAGATGGGGCTGAGG	240	DB	1261	GAGCCGAGTTCTGAGCTCTGTGCACTGGAGCCAGCCACAGCCGGCGGATGCATCTGCTG	1320
QY	763	TCACCTGTGGCCCAAGCCAGAAAGGGGAGGCGCACTCTCCCTCATTTCCCTTACAGCCCA	822	QY	1843	GGCAGTTTGTGGGAAAACCTATACCTCCGAGGCACTCTCCCTCATCACGGCTCGGACC	1902
DB	241	TCACCTGTGGCCCAAGCCAGAAAGGGGAGGCGCACTCTCCCTCATTTCCCTTACAGCCCA	300	DB	1321	GGCAGTTTGTGGGAAAACCTATACCTCCGAGGCACTCTCCCTCATCACGGCTCGGACC	1380
QY	823	AGTGAACCCCACTCTGGGGTCTCCAGCCAAACCACTCCACCGCAGTGTCTAATGCCCTGG	882	QY	1903	ACAGCTCTGCAGAACTCTCATTTCTTTTGGAGCAGGCACTGTGGGTAGAGTCTCTGGGG	1962
DB	301	AGTGAACCCCACTCTGGGGTCTCCAGCCAAACCACTCCACCGCAGTGTCTAATGCCCTGG	360	DB	1381	ACAGCTCTGCAGAACTCTCATTTCTTTTGGAGCAGGCACTGTGGGTAGAGTCTCTGGGG	1440
QY	883	ATCCATGAATTCGCGCGGGGTGCAACCCAGGGCTCAGAGAGAAGGTTTGGAGACAGCTG	942	QY	1963	TTCTCTGATGCAGAGGAAAGAAATATTTCTACAGATATTTTACAGATGAAGGCAAGCA	2022
DB	361	ATCCATGAATTCGCGCGGGGTGCAACCCAGGGCTCAGAGAGAAGGTTTGGAGACAGCTG	420	DB	1441	TTCTCTGATGCAGAGGAAAGAAATATTTCTACAGATATTTTACAGATGAAGGCAAGCA	1500
QY	943	CTTGACATCTGGAACCGCGCTGGAGAGAAATCTCTGCGCTCACTCCTCTACCAAGCTCTT	1002	QY	2023	ATTAGAGCCTTTAGGTTGGTCAAAATCAAAACAAAGAGCTCTGGGCCCTGTGTCTTGTGCC	2082
DB	421	CTTGACATCTGGAACCGCGCTGGAGAGAAATCTCTGCGCTCACTCCTCTACCAAGCTCTT	480	DB	1501	ATTAGAGCCTTTAGGTTGGTCAAAATCAAAACAAAGAGCTCTGGGCCCTGTGTCTTGTGCC	1560
QY	1003	CCAAGCTCCAGACCATCAGTCTCCAGCCAGGAGTCAACCAAGCCGCCACATCCACA	1062	QY	2083	TGGGTGTCTGCGCTGGCCCTGCACTTGGCCCTGATGCAGCAGATGAAGCGGAAGAAAACTC	2142
DB	481	CCAAGCTCCAGACCATCAGTCTCCAGCCAGGAGTCAACCAAGCCGCCACATCCACA	540	DB	1561	TGGGTGTCTGCGCTGGCCCTGCACTTGGCCCTGATGCAGCAGATGAAGCGGAAGAAAACTC	1620
QY	1063	GCAGTGTGGGAGCTGGGGATCCCCACCTCAGCCAGCCTAGCAACCCAGAGAGCAGGAG	1122	QY	2143	ACACTGACTTCAAGACAACCAACCAACCTCTGTCTACATTTACCTTGGCCAGGCTCTCDA	2202
DB	541	GCAGTGTGGGAGCTGGGGATCCCCACCTCAGCCAGCCTAGCAACCCAGAGAGCAGGAG	600	DB	1621	ACACTGACTTCAAGACAACCAACCAACCTCTGTCTACATTTACCTTGGCCAGGCTCTCDA	1680
QY	1123	GCTCTGGGACCAATGGCTCTGGATGAAAGTCAAGGAATTTACTACAGAAATCAGA	1182	QY	2203	GCTCAGCATTGGGACCCAGCTCAGAGACCTCTGCTCTCTGGCTCTGAGGGCATCTGG	2262
DB	601	GCTCTGGGACCAATGGCTCTGGATGAAAGTCAAGGAATTTACTACAGAAATCAGA	660	DB	1681	GCTCAGCATTGGGACCCAGCTCAGAGACCTCTGCTCTCTGGCTCTGAGGGCATCTGG	1740
QY	1183	GAAAGAGAGAGAGAAATCAGAGAAAGCAGGCCCCCAGTGGGAGCGTGTAGGAAGC	1242	QY	2263	CAAAAAAGACCTTTTTCAGTCCAGATGACCTCAGGAAGCATGGGTTCAGTGGGGCATC	2322
DB	661	GAAAGAGAGAGAGAAATCAGAGAAAGCAGGCCCCCAGTGGGAGCGTGTAGGAAGC	720	DB	1741	CAAAAAAGACCTTTTTCAGTCCAGATGACCTCAGGAAGCATGGGTTCAGTGGGGCATC	1800
QY	1243	CCCCACAGCGGCACACAGCCCTACAGCCCCACACACCCATGGGAGCCTTCTGTGAGA	1302	QY	2323	ATCTCCACTTCTCAAGATGGGTATTTCTCAAGAGCACCCCATCCCTCTGAGCTACAGC	2382
DB	721	CCCCACAGCGGCACACAGCCCTACAGCCCCACACACCCATGGGAGCCTTCTGTGAGA	780	DB	1801	ATCTCCACTTCTCAAGATGGGTATTTCTCAAGAGCACCCCATCCCTCTGAGCTACAGC	1860
QY	1303	GAGAGCTCTGTTCACATGGCCCTGGAAAAATGAGGAATTTTAAACAAAAATTCACACAG	1362	QY	2383	TTCAATTCACCTGTCTTCCAGAGTCTTTCAGCAATGTCTCTATCTCTGGAGATGAG	2442
DB	781	GAGAGCTCTGTTCACATGGCCCTGGAAAAATGAGGAATTTTAAACAAAAATTCACACAG	840	DB	1861	TTCAATTCACCTGTCTTCCAGAGTCTTTCAGCAATGTCTCTATCTCTGGAGATGAG	1920
QY	1363	CTGCTACTTCTACAAGACTCACCAGAGCCAGATCCCTGTCTCAAGAGAGCTGG	1422	QY	2443	AAGGGAGAGGTAAACATTTCTAATTTGCATCATAGATTTGGAAAAAGACGTAGAAGCATAT	2502
DB	841	CTGCTACTTCTACAAGACTCACCAGAGCCAGATCCCTGTCTCAAGAGAGCTGG	900	DB	1921	AAGGGAGAGGTAAACATTTCTAATTTGCATCATAGATTTGGAAAAAGACGTAGAAGCATAT	1980
QY	1423	CCTGATTATGTGGAGAGAAATCGAGGACATTTAATTTGAGATCAGAGACTTATTTGGCCCA	1482	QY	2503	GGAAATACATGGCCTGTTTGGGCACTCAACCAACGTTTCTATTTGGGCTGTTAAGTGAT	2040
DB	901	CCTGATTATGTGGAGAGAAATCGAGGACATTTAATTTGAGATCAGAGACTTATTTGGCCCA	960	DB	2041	GAGGGGAGAGAGATGGAGAACATCTTTCATCTCCGGCTGTCTCAGGGGAGAACCTTG	2100
QY	1483	GGCCTGGATACCAAGAACTCGCATAGTATCATCTCAGGGGCTCTCGAAATTTGGGAAG	1542	QY	2623	ATGCAAGTGGTCCCGTCCCTGAGCTGTCTGAGCCACACTCTCTGAGTCCCTCCAC	2682
DB	961	GGCCTGGATACCAAGAACTCGCATAGTATCATCTCAGGGGCTCTCGAAATTTGGGAAG	1020	DB	2101	ATGCAAGTGGTCCCGTCCCTGAGCTGTCTGAGCCACACTCTCTGAGTCCCTCCAC	2160
QY	1543	TCAACTGCGCAGGAGTGAAGGAAGCCTGGGGAGAGGCCAGCTGTATGGGACCGC	1602	QY	2683	TGCTTTGACGAGACTCGGAACAAAAAGTTTCTTGACACAGTGTATGGGCCCATTTTGAAGAA	2742
DB	1021	TCAACTGCGCAGGAGTGAAGGAAGCCTGGGGAGAGGCCAGCTGTATGGGACCGC	1080	DB	2161	TGCTTTGACGAGACTCGGAACAAAAAGTTTCTTGACACAGTGTATGGGCCCATTTTGAAGAA	2220
QY	1603	TTCCAGCATGTCTTTCTACTTTCAGCTCAGAGAGCTGGCCCGAGTCAAGGTGGTGTCTC	1662	QY	2743	ATGGGCACTGTGTAGAAAACAGACATGGAGCTTTAGTGTGCACTTTCTGCAATTAATTC	2802
DB	1081	TTCCAGCATGTCTTTCTACTTTCAGCTCAGAGAGCTGGCCCGAGTCAAGGTGGTGTCTC	1140	DB	2221	ATGGGCACTGTGTAGAAAACAGACATGGAGCTTTAGTGTGCACTTTCTGCAATTAATTC	2280
QY	1663	GCTGAGCTCATCGGAAAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCTGTCT	1722	QY	2803	AGCCGCCAGCTCAAGAGCTTTCAGCTGATTGAGGCGAGCAGCAGATCAACATGAGC	2862
DB	1141	GCTGAGCTCATCGGAAAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCTGTCT	1200	DB	2281	AGCCGCCAGCTCAAGAGCTTTCAGCTGATTGAGGCGAGCAGCAGATCAACATGAGC	2340
QY	1723	AGCCAGAGCGGCTCTTCTATCTCGATGATGATGAGCCAGATGGGTCTTGCAG	1782	QY	2863	CCACCATGGTAGTCTCTTTCAGTGGGTCCCAAGTCAAGATGCTTATTTGGCAGATTCTC	2922
DB	1201	AGCCAGAGCGGCTCTTCTATCTCGATGATGATGAGCCAGATGGGTCTTGCAG	1260				

Db 2341 CCACACAGTAGTCTCTGTTCCAGTGGTCCAGTCCAGATCCAGATTCCTC 2400
 Qy 2923 TTCTCGTCTCAAGTCAACCAAGCTGAGGAGCTGGAACCTAAGTGGAACTCGCTG 2982
 Db 2401 TTCTCGTCTCAAGTCAACCAAGCTGAGGAGCTGGAACCTAAGTGGAACTCGCTG 2460
 Qy 2983 AGCCACTCTGCAAGTGAAGTCTTTTGAAGACCTGAGAGCCCTCGCTGCTCCTGGAG 3042
 Db 2461 AGCCACTCTGCAAGTGAAGTCTTTTGAAGACCTGAGAGCCCTCGCTGCTCCTGGAG 2520
 Qy 3043 ACCCTGCGTGTGGCTGTGCTGCTCAAGCTGAGAGCTGCAAGGACCTTCCCTTTGGG 3102
 Db 2521 ACCCTGCGTGTGGCTGTGCTGCTCAAGCTGAGAGCTGCAAGGACCTTCCCTTTGGG 2580
 Qy 3103 CTGAGAGCCAAACAGACCTGAGCCGAGCTGGAAGCTTCAATGCTGCTGAGATGCT 3162
 Db 2581 CTGAGAGCCAAACAGACCTGAGCCGAGCTGGAAGCTTCAATGCTGCTGAGATGCT 2640
 Qy 3163 GGAGCCAAACACCTTTGCCAGAGCTGAGACGAGCTGCAAGCTACAGGAGCTGCAAG 3222
 Db 2641 GGAGCCAAACACCTTTGCCAGAGCTGAGACGAGCTGCAAGCTACAGGAGCTGCAAG 2700
 Qy 3223 CTGCTCAGCTGTGGCTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3282
 Db 2701 CTGCTCAGCTGTGGCTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760
 Qy 3283 AGCCCCAGCTGAGAGCTGAGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 3342
 Db 2761 AGCCCCAGCTGAGAGCTGAGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2820
 Qy 3343 CTGCTCTGTGAGGGCTCAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3402
 Db 2821 CTGCTCTGTGAGGGCTCAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2869
 Qy 3403 ACAACTCTGAGTATGAGATGAGCAGGAGAACTGAGGGCCCTGAGCAGGAGAACTCTCAG 3462
 Db 2870 ----- 2869
 Qy 3463 CTGCTCATCTTCCAGCAGCAGGAAACCAAGTGTGATGACCCCTACTGAGGGCTGATAGC 3522
 Db 2870 -----GGAACCAAGTGTGATGACCCCTACTGAGGGCTGATAGC 2910
 Qy 3523 GGAGAGTCAAGTAATAGCATCTCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGGCG 3582
 Db 2911 GGAGAGTCAAGTAATAGCATCTCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGGCG 2970
 Qy 3583 GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCTTGGAGCTGAGCAAGATCTTCCCAATT 3642
 Db 2971 GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCTTGGAGCTGAGCAAGATCTTCCCAATT 3030
 Qy 3643 GCTGAGATTGCAAGGAAAGCTCCAGAGGTAGTACCGGTGGAACCTTTGCGGTGCT 3702
 Db 3031 GCTGAGATTGCAAGGAAAGCTCCAGAGGTAGTACCGGTGGAACCTTTGCGGTGCT 3090
 Qy 3703 TCTCTGCTCTCAAGGGACCTGCAATAGAGCCCTTGGGAGCTGACGATCTTCTGG 3762
 Db 3091 TCTCTGCTCTCAAGGGACCTGCAATAGAGCCCTTGGGAGCTGACGATCTTCTGG 3150
 Qy 3763 GGCCCCCAGGGCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTTGACCGAGTT 3822
 Db 3151 GGCCCTGAGGAAATGTGGATGTTGAGTTGATTGATAGAGCAGACAGATACAGCGTT 3210
 Qy 3823 CACTTCCCTGTAGTGGCTCTCAACCGCTGCGCCCAACACCGGTCTCTGCTTTGTGATGAGA 3882
 Db 3211 TGGTTCCCACTGCTGGCTGATCTGTGCTGAGCCAGCCAGCCGCTCGGCTTCTGGTAAG 3270
 Qy 3883 GAAGCGGTACCGTTGAGATTGAATTTCTGCTGTTGGGACAGTTCTCTGG ---GTGAGATC 3939
 Db 3271 GATGAGTCAAGTGAAGTTGGTTTGGTTTCTGAGTCAAGCTGCGCTGAGACCTG 3330
 Qy 3940 AACCCACAGCAGCTGAGTGGTGGCAGGGCCCTCTGCTGAGATCAAGGCTGAGCCTGGA 3999

Db 3331 CAGCACCATGAACAGTGGCTGGTGGGGGGCCCTTGTGTTGATGTACTGACAGAGCCAGAG 3390
 Qy 4000 GCTGTGGAAGCTGTGACCTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTC 4059
 Db 3391 GAGGCTGTGCGCGAAATCCACCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTC 3450
 Qy 4060 TCCCTGTTTCCAAATGGCCCACTTTAAAGAGGAGGGGATGCTCTCTGGAGAAAGCCAGCCAGG 4119
 Db 3451 TCCTGTTTCTCGTTGCCCCATTTAAAGAAATGAAGGGATGGTCTCTGGAGCATCCAGCCCGG 3510
 Qy 4120 GTGAGCTGATCATCATAGTTCTGSAAGAACCCAGCTTCTCCCTTGGAGTCTCTCTG 4179
 Db 3511 GTGGAGCCTTTCTATGCTGCTCTGSAAGAACCCAGCTTCTCTGATGGGATCTCTGCTG 3570
 Qy 4180 AAATGATCCATTAATGCTGCTTCAATCCCGTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTG 4239
 Db 3571 CGGATCGCAGTGGAGCTGCTCTCCATCCCTCACTTCCACACATGATCTATTTAT 3630
 Qy 4240 CGCTGCTCACTCTGAGGAGTCACTTCCCTCACTCTACCTGATCCCAAGTCACTGCTTCAAT 4299
 Db 3631 CACCCCAACCCCGAAGATATTAAAGTTCCACTTGTACCTTGTCCCGAGCGCTTGTCTA 3690
 Qy 4300 CGAAGGAAGCTGGA 4313
 Db 3691 ACAAGGCGATAGA 3704
 RESULT 11
 AAA78392
 ID AAA78392 standard; cDNA; 2657 BP.
 XX
 AC AAA78392;
 XX
 DT 20-NOV-2000 (first entry)
 XX
 DE Human secreted protein gene 12 SEQ ID NO:22.
 XX
 KW Human; secreted protein; cytostatic; antineoplastic; antidiabetic;
 KW antiinflammatory; ophthalmological; antirheumatic; antiarthritic;
 KW antiproliferative; angiogenic; cardiac; anti-HIV; neurotrophic;
 KW immunoregulatory; antimicrobial; antiparkinsonian; cancer;
 KW immune system disorder; angiogenesis; hyperproliferative disorder;
 KW cardiovascular disorder; apoptosis; neurological disease;
 KW infectious disease; wound healing; ss.
 OS Homo sapiens.
 XX
 XX WO200035937-A1.
 PN
 XX
 PD 22-JUN-2000.
 XX
 PF 16-DEC-1999; 99WO-US29950.
 XX
 PR 17-DEC-1998; 98US-0112809.
 PR 18-DEC-1998; 98US-0113006.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Ebner R, Rosen CA, Endress GA, Soppet DR, Ni J;
 PI Duan DR, Moore PA, Shi Y, Lafleur DW, Olsen HS, Florence K;
 XX
 DR WPI; 2000-431566/37.
 DR P-PSDB; AAB24448.
 XX
 PT Forty seven human nucleic acids encoding secreted proteins, useful in
 PT the treatment, prevention and diagnosis of cancers, disorders of the
 PT immune system, angiogenesis disorders, neurological diseases and
 PT hyperproliferative disorders -
 XX
 PS Claim 1; Page 449-450; 562pp; English.
 XX
 CC The polynucleotide sequence given in AAA78381 to AAA78432 encode the
 CC human secreted proteins given in AAB24437 to AAB24604. Human secreted

proteins have activities based on the tissues and cells the genes are expressed in. Examples of activities include: cytostatic; antianaemic; antiadipatic; antiinflammatory; ophthalmological; anti rheumatic; antiarthritic; anipsostatic; antiangiogenic; cardiant; anti-HIV; nootropic; neuroprotective; antimicrobial and antiparkinsonian. Human secreted protein polynucleotides, polypeptides and/or agonists may be useful in treating, preventing, and/or diagnosing other diseases, disorders, and/or conditions such as: (a) cancers; (b) disorders of the immune system; (c) angiogenesis disorders; (d) hyperproliferative disorders; (e) cardiovascular disorders; (f) diseases associated with increase apoptosis; (g) neurological diseases; and (h) infectious diseases. They are also used to promote wound healing. AAA78372 to AAA78380 and AAA24436 represent sequences used in the exemplification of the present invention.

Sequence 2657 BP: 664 A; 696 C; 726 G; 571 T; 0 other;

Query Match 47.3%; Score 2572.6; DB 21; Length 2657;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 2614; Conservative 0; Mismatches 14; Indels 3;

2814	QAAGAAAGCTTTACGTGATTTGAGGCGCAGCGCAGACACAGATCAAACTGAGAGCCCCACCATGGT	2873
QY		
9	GAAGAAGCTTTACGTGATTTGAGGCGCAGCGCAGACACAGATCAAACTGAGAGCCCCACCATGGT	68
Db		
2874	AGTCCCTGTTACAGGTGGGTCCCAAGTTCACAGATGCCTATTGGCAGATTCTCTTCCTCGGCTT	2933
QY		
69	AGTCCCTGTTACAGGTGGGTCCCAAGTTCACAGATGCCTATTGGCAGATTCTCTTCCTCGGCTT	128
Db		
2934	CAAGGTCACACAGAAACCTGAAGGAGCTGGACCTTAAGTGGAAACTCGCTGAGCCACTCTGCG	2993
QY		
129	CAAGGTCACACAGAAACCTGAAGGAGCTGGACCTTAAGTGGAAACTCGCTGAGCCACTCTGCG	188
Db		
2994	AGTGAAGAGCTTTGTAAAGACCTTGAGACCGCCTCGCTGCGCTCTGAGAGACCTTGCGGGTT	3053
QY		
189	AGTGAAGAGCTTTGTAAAGACCTTGAGACCGCCTCGCTGCGCTCTGAGAGACCTTGCGGGTT	248
Db		
3054	GGCTGGGCTGTGGCTCACAGCTGAGGAGCTGCAAGGACCTTGCCCTTGGGCTGAGAGCCAA	3113
QY		
249	GGCTGGGCTGTGGGCTCAAGCTGAGGACTGCAAGGACCTTGCCCTTGGGCTGAGAGCCAA	308
Db		
3114	CCAGACCCCTGAACGAGCTGGAACCTGAGCTTCAATGTGCTCAACGATGCTGAGAGCCAAACA	3173
QY		
309	CCAGACCCCTGACCGAGCTGAGCTGAGCTTCAATGTGCTCAACGATGCTGAGAGCCAAACA	368
Db		
3174	CCCTTGGCAGAGACTGAGACAGCCGAGCTGACGCACTGAGAGCTGCTGACGTG	3233
QY		
369	CCCTTGGCAGAGACTGAGACAGCCGAGCTGACGCACTGAGAGCTGCTGACGTG	428
Db		
3234	TGGGCTCACGCTCAGCTGCTGCCAGGACCTGGGCTCTGTGCTTAAGTGCAGAGCCCAAGCT	3293
QY		
429	TGGGCTCACGCTCAGCTGCTGCCAGGACCTGGGCTCTGTGCTTAAGTGCAGAGCCCAAGCT	488
Db		
3294	GAAGGAGCTAGACCTGACAGCAGAAACAACTGGATGACGTTGGCGTGCAGCTGCTGTGA	3353
QY		
489	GAAGGAGCTAGACCTGACAGCAGAAACAACTGGATGACGTTGGCGTGCAGCTGCTGTGA	548
Db		
3354	GGGGCTCAGGCATCCTCGCTGCAAACTCATACGCTGGGCTGAGACAGCAAACTGTGAG	3413
QY		
549	GGGGCTCA-GCATCTGCTGCAAACTCATACGCTGGGCTGAGACAGCAAACTGTGAG	607
Db		
3414	TGATGAGATGAGGACGAGAACTGAGGGGCTGGAGCAGAGAAAACCTCAGCTGCTCATCTT	3473
QY		
608	TGATGAGATGAGGACGAGAACTGAGGGGCTGGAGCAGAGAAAACCTCAGCTGCTCATCTT	667
Db		
3474	CAGCAGACGGAACCAAGTGTGATGACCCCTACTCAGGGCCTGGATACGGAGAGATGAG	3533
QY		
668	CAGCAGACGGAACCAAGTGTGATGACCCCTACTGAGGGCCTGGATACGGAGAGATGAG	726
Db		
3534	TNATAGACATCTCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGGGGCTTCCCATGT	3593
QY		
727	TNATAGACATCTCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGGGGCTTCCCATGT	786
Db		

3594	Qy	TGCTCAGGCTTAATCTCTCAAACTCTCTGAGCTGAGCAAGATCTTCCCAAATTTGCTCTGAGATTGC	3655
787	Db	TGCTCAGGCTTAATCTCTCAAACTCTCTGAGCTGAGCAAGATCTTCCCAAATTTGCTCTGAGATTGC	846
3654	Qy	AGAGGAAAGCTCCCCAGAGGTAGTACCGGTGGAACTCTTTGTGGTGCTTCTCTCTGCTCTC	3713
847	Db	AGAGGAAAGCTCCCCAGAGGTAGTACCGGTGGAACTCTTTGTGGTGATGCCCTCTCTCTGCTCTC	906
3714	Qy	TCAAGGGGACCTTGCAATACGAAAGCCTTTGGGGACTGACGATGACTTTCTGGGGCCCCACGGG	3773
907	Db	TCAAGGGGACCTTGCAATACGAAAGCCTTTGGGGACTGACGATGACTTTCTGGGGCCCCACGGG	966
3774	Qy	GCCCTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTTGTAACGAGTTTCACATTCCTCTGT	3833
967	Db	GCCCTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTTGTAACGAGTTTCACATTCCTCTGT	1026
3834	Qy	AGCTGGCTCTACCGCTGGCCAAACCGGTCTCTGCTTTGTGATGAGAGAACGGGTGAC	3893
1027	Db	AGCTGGCTCTACCGCTGGCCAAACCGGTCTCTGCTTTGTGATGAGAGAACGGGTGAC	1086
3894	Qy	CGTTGAGATTGAATTTCTGTGTGGGACCAAGTTCTTGGGTGAGATCAACCCACAGCACAG	3953
1087	Db	CGTTGAGATTGAATTTCTGTGTGGGACCAAGTTCTTGGGTGAGATCAACCCACAGCACAG	1146
3954	Qy	CTGGATTGTGGCAGGGCCTCTGCTGAGACATCAAGGCTGAGCTCTGGAGCTGTGGAGCTGT	4013
1147	Db	CTGGATTGTGGCAGGGCCTCTGCTGAGACATCAAGGCTGAGCTCTGGAGCTGTGGAGCTGT	1206
4014	Qy	GCACTTCCTCACTTTGTGGCTCTCAAGGGGGCCATGTGGACACATCCCTGTTCCAAAT	4073
1207	Db	GCACTTCCTCACTTTGTGGCTCTCAAGGGGGCCATGTGGACACATCCCTGTTCCAAAT	1266
4074	Qy	GGCCCACTTTAAAGAGAGGGGATGCTCTCTGGAGAAAGCCACGAGGTGGAGCTGCATCA	4133
1267	Db	GGCCCACTTTAAAGAGAGGGGATGCTCTCTGGAGAAAGCCACGAGGTGGAGCTGCATCA	1326
4134	Qy	CATAGTTCTGGAAAAACCCAGCTTTCTCCCTTTGGGAGTCCCTCTGAAAAATGATCATAA	4193
1327	Db	CATAGTTCTGGAAAAACCCAGCTTTCTCCCTTTGGGAGTCCCTCTGAAAAATGATCATAA	1386
4194	Qy	TGCCCCCTGCTTCACTTCCCGTACACTCTGTGGTGTGCTTTTACACCGCGTCCATCCTGA	4253
1387	Db	TGCCCCCTGCTTCACTTCCCGTACACTCTGTGGTGTGCTTTTACACCGCGTCCATCCTGA	1446
4254	Qy	GGAAGTCACCTTCCACTCTACCTGATCCCAAATGACTGCTCCATTCGGAAGGAACTGGGA	4313
1447	Db	GGAAGTCACCTTCCACTCTACCTGATCCCAAATGACTGCTCCATTCGGAAGGAACTGGGA	1506
4314	Qy	GCTCTGTATCGAAGGCCCTGGAGAAACCAAGCTGTTCTCGGAGTTCTACGTTGGCCACTT	4373
1507	Db	GCTCTGTATCGAAGGCCCTGGAGAAACCAAGCTGTTCTCGGAGTTCTACGTTGGCCACTT	1566
4374	Qy	GGATCAGGGATCAGGCTGCAAGTGAAGACAAGAGATGAGACTCTGCTGTGGGAGGC	4433
1567	Db	GGATCAGGGATCAGGCTGCAAGTGAAGACAAGAGATGAGACTCTGCTGTGGGAGGC	1626
4434	Qy	CTTGTGTGAACACAGGAGATCTCATGCTCGTCAAATACCTCTGATCCCTCCAGCCGCATAGC	4493
1627	Db	CTTGTGTGAACACAGGAGATCTCATGCTCGTCAAATACCTCTGATCCCTCCAGCCGCATATC	1686
4494	Qy	CGTACCTTCACTCTGGATCCCCCGACAGTTGCTGCACTTTGTGGACAGATTCGAGAGCA	4553
1687	Db	CGTACCTTCACTCTGGATCCCCCGAGTTGCTGCACTTTGTGGACAGATTCGAGAGCA	1746
4554	Qy	GCTGATAGCCGAGTGACATCGGTGGAGGTTGCTTTGGACAAACTGCATGGACAGGTGCT	4613
1747	Db	GCTGATAGCCGAGTGACATCGGTGGAGGTTGCTTTGGACAAACTGCATGGACAGGTGCT	1806
4614	Qy	GAGCCAGGACGATACGAGAGGTTGCTGGCTGAGAACACGAGGCCACAGCAGATGCGGAA	4673
1807	Db	GAGCCAGGACGATACGAGAGGTTGCTGGCTGAGAACACGAGGCCACAGCAGATGCGGAA	1866
4674	Qy	GCTGTTTCAAGCTTGAGCCAGTCTCTGGGACCGGAAAGTGCAAGATGGAATCTCTACGAAGCCCT	4733

Db	1867	GCTGTTTCAGCTTGAGCCAGTCTCTGGGACGCGAAGTGCACAAAGATGGACTCTACCAAGCCCT	1926	PR	31-JAN-2000;	2000US-0179065.
Qy	4734	GAAGGAGAGCCATCCTCAGCTCATCTGGAATGAACTCTGGGAGAGGGGAGCAAAAAGGACT	4793	PR	04-FEB-2000;	2000US-0180628.
Db	1927	GAAGGAGAGCCATCCTCAGCTCATCTGGAATGAACTCTGGGAGAGGGGAGCAAAAAGGACT	1985	PR	24-FEB-2000;	2000US-0184664.
Qy	4794	CCTGCCACTCAGCAGCTGAAGTATCAACACCGCCCTTGACCCCTTGAGTCTGGCTTTGG	4853	PR	02-MAR-2000;	2000US-0186350.
Db	1986	CCTGCCACTCAGCAGCTGAAGTATCAACACCGCCCTTGAGTCTGGCTTTGG	2045	PR	16-MAR-2000;	2000US-0189874.
Qy	4854	CTGACCCCTTTTGGGTCTCAGTTTCTTCTCTGCAAAAGTGGCATCTGTTGGCT	4913	PR	17-MAR-2000;	2000US-0190076.
Db	2046	CTGACCCCTTTTGGGTCTCAGTTTCTTCTCTGCAAAAGTGGCATCTGTTGGCT	2105	PR	18-APR-2000;	2000US-0198123.
Qy	4914	TCAGACATTAAGTAAATGGAATCTTGATGATGCTTTGCTGGGCATTTATGTGTCATGCC	4973	PR	19-MAY-2000;	2000US-0205515.
Db	2106	TCAGACATTAAGTAAATGGAATCTTGATGATGCTTTGCTGGGCATTTATGTGTCATGCC	2165	PR	07-JUN-2000;	2000US-0209467.
Qy	4974	AGGGATGCCACAGGGGGCCCCAGTCCAGGTGGCCTAACACGATCTCAGGGAATGTCCATC	5033	PR	28-JUN-2000;	2000US-0214886.
Db	2166	AGGGATGCCACAGGGGGCCCCAGTCCAGGTGGCCTAACACGATCTCAGGGAATGTCCATC	2225	PR	30-JUN-2000;	2000US-0215135.
Qy	5034	TGAGCTGGCAAGACCCCTCGACACCTCATAGAGCCTCATCTGGTGGCCACAGCAGCCAA	5093	PR	07-JUL-2000;	2000US-0216647.
Db	2226	TGAGCTGGCAAGACCCCTCGACACCTCATAGAGCCTCATCTGGTGGCCACAGCAGCA	2285	PR	07-JUL-2000;	2000US-0216680.
Qy	5094	GCTTAGAGCCCTCCGATCCCATCCAGGCGCAAGAGGAATAGGAGGACATGGAAACCAT	5153	PR	11-JUL-2000;	2000US-0217487.
Db	2286	AGCCTAGAGCCCTCCGATCCCATCCAGGCGCAAGAGGAATAGGAGGACATGGAAACCAT	2345	PR	11-JUL-2000;	2000US-0217496.
Qy	5154	TTGCCCTCTGGCTGTGTACAGGCTGAGCCCAAAATTTGGGTTCAGCGTGGGAGGCCACG	5213	PR	14-JUL-2000;	2000US-0218290.
Db	2346	TTGCCCTCTGGCTGTGTACAGGCTGAGCCCAAAATTTGGGTTCAGCGTGGGAGGCCACG	2405	PR	26-JUL-2000;	2000US-0220963.
Qy	5214	TGGATTCTTTGGCTTTGTACAGGAAGATCTACAGGACCAAGCAAGCAAGCAAGTAAAGTGAAG	5273	PR	26-JUL-2000;	2000US-0220964.
Db	2406	TGGATTCTTTGGCTTTGTACAGGAAGATCTACAGGACCAAGCAAGCAAGTAAAGTGAAG	2465	PR	14-AUG-2000;	2000US-0224518.
Qy	5274	GAAGTTTATTCAGAAAATAAAGAGTATCACAGCTCTTTTAGAAATTTGTCTAGCAGGCTT	5333	PR	14-AUG-2000;	2000US-0224519.
Db	2466	GAAGTTTATTCAGAAAATAAAGAGTATCACAGCTCTTTTAGAAATTTGTCTAGCAGGCTT	2525	PR	14-AUG-2000;	2000US-0225213.
Qy	5334	TCAGTTTATTCAGAAAACCCCTATAAATTTAAATTTTAAATTTTAAATTTTAAATTTA	5393	PR	14-AUG-2000;	2000US-0225214.
Db	2526	TCAGTTTATTCAGAAAACCCCTATAAATTTAAATTTTAAATTTTAAATTTTAAATTTA	2585	PR	22-AUG-2000;	2000US-0226681.
Qy	5394	AAAAATACAAAAAGAAAAATGAAAAATGAAAAATGAAAAATGAAAAATGAAAAATG	5444	PR	22-AUG-2000;	2000US-0226688.
Db	2586	AAAAATACAAAAAGAAAAATGAAAAATGAAAAATGAAAAATGAAAAATGAAAAATG	2636	PR	22-AUG-2000;	2000US-0227182.

RESULT 12
AAL01744
ID AAL01744 standard; cdna; 578 BP.
XX AAL01744;
AC AAL01744;
XX
XX 21-NOV-2001 (first entry)
DE Human reproductive system related antigen cdna SEQ ID NO: 1745.
DE Human reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ss.
XX Homo sapiens.
XX WO20015320-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01339.
XX

PR	02-OCT-2000;	2000US-0237037.	XX	Claim 1; SEQ ID NO 1745; 1297pp + Sequence Listing; English.
PR	02-OCT-2000;	2000US-0237038.	PS	
PR	02-OCT-2000;	2000US-0237039.	XX	
PR	02-OCT-2000;	2000US-0237040.	CC	The present invention provides the protein and coding sequences of a
PR	13-OCT-2000;	2000US-0239935.	CC	number of human reproductive system related antigens. These can be used
PR	13-OCT-2000;	2000US-0239937.	CC	in the prevention and treatment of reproductive system disorders,
PR	20-OCT-2000;	2000US-0240960.	CC	including cancer. The present sequence is a coding sequence of the
PR	20-OCT-2000;	2000US-0241221.	CC	invention.
PR	20-OCT-2000;	2000US-0241785.	XX	
PR	20-OCT-2000;	2000US-0241786.	XX	
PR	20-OCT-2000;	2000US-0241787.	XX	
PR	20-OCT-2000;	2000US-0241808.	XX	
PR	20-OCT-2000;	2000US-0241809.	XX	
PR	20-OCT-2000;	2000US-0241826.	XX	
PR	01-NOV-2000;	2000US-0244617.	XX	
PR	08-NOV-2000;	2000US-0246474.	XX	
PR	08-NOV-2000;	2000US-0246475.	XX	
PR	08-NOV-2000;	2000US-0246476.	XX	
PR	08-NOV-2000;	2000US-0246477.	XX	
PR	08-NOV-2000;	2000US-0246478.	XX	
PR	08-NOV-2000;	2000US-0246523.	XX	
PR	08-NOV-2000;	2000US-0246524.	XX	
PR	08-NOV-2000;	2000US-0246525.	XX	
PR	08-NOV-2000;	2000US-0246526.	XX	
PR	08-NOV-2000;	2000US-0246527.	XX	
PR	08-NOV-2000;	2000US-0246528.	XX	
PR	08-NOV-2000;	2000US-0246532.	XX	
PR	08-NOV-2000;	2000US-0246609.	XX	
PR	08-NOV-2000;	2000US-0246610.	XX	
PR	08-NOV-2000;	2000US-0246611.	XX	
PR	08-NOV-2000;	2000US-0246613.	XX	
PR	17-NOV-2000;	2000US-0249207.	XX	
PR	17-NOV-2000;	2000US-0249208.	XX	
PR	17-NOV-2000;	2000US-0249209.	XX	
PR	17-NOV-2000;	2000US-0249210.	XX	
PR	17-NOV-2000;	2000US-0249211.	XX	
PR	17-NOV-2000;	2000US-0249212.	XX	
PR	17-NOV-2000;	2000US-0249213.	XX	
PR	17-NOV-2000;	2000US-0249214.	XX	
PR	17-NOV-2000;	2000US-0249215.	XX	
PR	17-NOV-2000;	2000US-0249216.	XX	
PR	17-NOV-2000;	2000US-0249217.	XX	
PR	17-NOV-2000;	2000US-0249218.	XX	
PR	17-NOV-2000;	2000US-0249244.	XX	
PR	17-NOV-2000;	2000US-0249245.	XX	
PR	17-NOV-2000;	2000US-0249264.	XX	
PR	17-NOV-2000;	2000US-0249265.	XX	
PR	17-NOV-2000;	2000US-0249297.	XX	
PR	17-NOV-2000;	2000US-0249299.	XX	
PR	17-NOV-2000;	2000US-0249299.	XX	
PR	01-DEC-2000;	2000US-0250160.	XX	
PR	01-DEC-2000;	2000US-0250391.	XX	
PR	05-DEC-2000;	2000US-0251030.	XX	
PR	05-DEC-2000;	2000US-0251988.	XX	
PR	05-DEC-2000;	2000US-0256719.	XX	
PR	06-DEC-2000;	2000US-0251479.	XX	
PR	08-DEC-2000;	2000US-0251856.	XX	
PR	08-DEC-2000;	2000US-0251868.	XX	
PR	08-DEC-2000;	2000US-0251869.	XX	
PR	08-DEC-2000;	2000US-0251989.	XX	
PR	08-DEC-2000;	2000US-0251990.	XX	
PR	11-DEC-2000;	2000US-0254097.	XX	
PR	05-JAN-2001;	2001US-0259678.	XX	
XX	(HUMA-) HUMAN GENOME SCI INC.		XX	
PA	Rosen CA, Barash SC, Ruben SM;		XX	
PI	WPI; 2001-465570/50.		XX	
DR	F-PSDB; AAM95774.		XX	
DR	Isolated nucleic acid molecule encoding a reproductive system antigen		XX	
PT	is used in preventing, treating or ameliorating a medical condition -		PT	

Query Match 8.9%; Score 484.4; DB 22; Length 578;
Best Local Similarity 99.4%; Pred. No. 5e-110; Mismatches 2; Indels 1; Gaps 1;
Matches 496; Conservative 0;

QY 1375 CAAGACCTCACCCAGAGCCCAAGATCCCTGGTCAAGAGAAAGCTGGCTGATTATGTG 1434
DB 2 CAAGACCTCACCCAGAGCCCAAGATCCCTGGTCAAGAGAAAGCTGGCTGATTATGTG 61

QY 1435 GAGGAGATCGAGGACATTTAATTGAGATCAGAGACTTATTTGGCCAGGCTGGATACC 1494
DB 62 GAGGAGATCGAGGACATTTAATTGAGATCAGAGACTTATTTGGCCAGGCTGGATACC 121

QY 1495 CAAGAACCTCGCATAGTATCTACTGCAGGGGGCTGCTGGAAATTGGGAAGTCAACACTGGCC 1554
DB 122 CAAGAACCTCGCATAGTATCTACTGCAGGGGGCTGCTGGAAATTGGGAAGTCAACACTGGCC 181

QY 1555 AGGCAGGTGAAGGAAGCCCTGGGGGAGAGCCAGCTGTATGGGAGCCGCTTCAGCATGTC 1614
DB 182 AGGCAGGTGAAGGAAGCCCTGGGGGAGAGCCAGCTGTATGGGAGCCGCTTCAGCATGTC 241

QY 1615 TTCTACTTCAGCTGCAGAGAGCTGGCCAGTCCAAAGTGGTGGTCTCGCTGAGCTCATC 1674
DB 242 TTCTACTTCAGCTGCAGAGAGCTGGCCAGTCCAAAGTGGTGGTCTCGCTGAGCTCATC 301

QY 1675 GGAAAGATGGACAGCCACTCCGGCTCCATTAGACAGATCCTGTCTAGGCCAGAGCGG 1734
DB 302 GGAAAGATGGACAGCCACTCCGGCTCCATTAGACAGATCCTGTCTAGGCCAGAGCGG 361

QY 1735 CTGCTCTTCATCCTCGATGGTGTAGATGAGCAGATGGTCTTGACAGAGCCGAGTTCT 1794
DB 362 CTGCTCTTCATCCTCGATGGTGTAGATGAGCAGATGGTCTTGACAGAGCCGAGTTCT 421

QY 1795 GAGCTCTGTCTGCACTGGAGCCAGCCACAG-CCGGCGGATGCACCTCTGGGAGTTTGTCT 1853
DB 422 GAGCTCTGTCTGCACTGGAGCCAGCCACAGCCCGCGGATGCACCTCTGGGAGTTTGTCT 481

QY 1854 GGGGAAAACATATCTTCCC 1872
DB 482 GGGGAAAACATATCTTCCC 500

RESULT 13
ABL97037
ID ABL97037 standard; cDNA; 578 BP.
XX ABL97037;
AC ABL97037;
XX ABL97037;
DT 21-JUN-2002 (first entry)
XX Human testicular antigen encoding cDNA SEQ ID NO: 705.
DE Human testicular antigen; testes; cancer; metastasis; immune disorder;
XX Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW reproductive system disorder; urinary system disorder; gene therapy;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disease; infection; cytostatic; gene; ss.
XX Homo sapiens.
OS Homo sapiens.
XX WO200155317-A2.
FN WO200155317-A2.
XX 02-AUG-2001.
PD 02-AUG-2001.
XX

PF 17-JAN-2001; 2001WO-US01329.
 XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226688.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.

PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Barash SC, Ruben SM;
 PI WPI; 2001-483232/52.
 XX Nucleic acids encoding 973 human testicular antigen polypeptides,
 PT

XX	30-JAN-2002; 2002WO-US02781.
XX	
XX	30-JAN-2001; 2001US-265305P.
PR	31-JAN-2001; 2001US-265682P.
PR	09-FEB-2001; 2001US-267568P.
PR	21-MAR-2001; 2001US-278651P.
PR	28-APR-2001; 2001US-287112P.
PR	16-MAY-2001; 2001US-291631P.
PR	12-JUL-2001; 2001US-305484P.
PR	20-AUG-2001; 2001US-313999P.
PR	27-NOV-2001; 2001US-333626P.
XX	(CORI-) CORIXA CORP.
XX	Benson DR, Kalos MD, Lodes MJ, Persing DH, Hepler WT, Jiang Y;
XX	WPI; 2002-627435/67.
DR	
PT	New isolated polynucleotide and pancreatic tumor polypeptides, useful
PT	for diagnosing, preventing and/or treating cancer, particularly
PT	pancreatic cancer
PS	
XX	Claim 1; SEQ ID NO 3298; 300pp + Sequence Listing; English.
XX	
CC	The invention relates to an isolated polynucleotide (I) comprising: (a)
CC	any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145);
CC	(b) complements of (a); (c) sequences consisting of at least 20
CC	contiguous residues of (a); (d) sequences that hybridize to (a), under
CC	moderately stringent conditions; (e) sequences having at least 75% or 90%
CC	identity to (a); or (f) degenerate variants of (a). Polypeptides
CC	(ABP68596-ABP68637) encoded by (I) and oligonucleotide can be used to
CC	detect cancer in a patient and compositions comprising polypeptides,
CC	polynucleotides, antibodies, fusion proteins, T cell populations and
CC	antigen presenting cells expressing the polypeptide are useful in
CC	treating pancreatic cancer and stimulating an immune response. The
CC	polynucleotides can be used as probes or primers for nucleic acid
CC	hybridisation, in the design and preparation of ribozyme molecules for
CC	inhibiting expression of the tumour polypeptides and proteins in the
CC	tumour cells, in vaccines and for gene therapy.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	
SQ	Sequence 487 BP; 114 A; 130 C; 133 G; 109 T; 1 other;
Query Match 8.9%; Score 482.8; DB 24; Length 487;	
Best Local Similarity 99.4%; Pred. No. 1.1e-109;	
Matches 484; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
QY	4744 CATCCTCACCTCATTTGGAACCTCTGGAGAGCGGCAGCAAAAGGCCACTCTGCCACTC 4803
Db	
	487 CATCCTCACCTTATTATGNAACCTCTGGAGAGCGGCAGCAAAAGGCCACTCTGCCACTC 428
QY	4804 AGCAGCTGAAGTATCAACACAGCCCTTGACCCTTCAGTCTGTGGCTTTGGCTGACCCCTTC 4863
Db	427 AGCAGCTGAAGTATCAACACAGCCCTTGACCCTTCAGTCTGTGGCTTTGGCTGACCCCTTC 368
QY	4864 TTGGGGTCTCAGTTTTCTTTTCTCTGCAACAAGTTGCCATCTGGTTTGGCTTCCAGCACTA 4923
Db	367 TTTGGGTCTCAGTTTTCTTTTCTCTGCAACAAGTTGCCATCTGGTTTGGCTTCCAGCACTA 308
QY	4924 AAGTAATGGACCTTTGATGATGCTTTTGCTGGGCATTTATGTGTCATGCGCCGGATGCCA 4983
Db	307 ARGTAATGGAACCTTTGATGATGCTTTTGCTGGGCATTTATGTGTCATGCGCCGGATGCCA 248
QY	4984 CAGGGGGCCCCAGTCCAGGTGGCCCTAACAGCATCTCAGGGAATGTCCATCTGGAGCTGGC 5043
Db	247 CAGGGGGCCCCAGTCCAGGTGGCATACAGCATCTCAGGGAATGTCCATCTGGAGCTGGC 188
QY	5044 AAGAACCCTTCAGACCTTCATAGACCTTCATCTGGTGGCCACAGCAGCCAGCCCTAGAGCC 5103
Db	187 AAGACCCCTTCAGACCTTCATAGACCTTCATCTGGTGGCCACAGCAGCCAGCCCTAGAGCC 128

[illegible]

QY 5104 CTCGCGATCCATCCAGGCGCAAGAGGAATAGGAGGACATGGACCATTTGGCTCTCG 5163
 Db 127 CTCGCGATCCATCCAGGCGCAAGAGGAATAGGAGGACATGGACCATTTGGCTCTCG 68
 QY 5164 CTGTGTACAGGCTGAGCCGCCAAAATTGGGGTTTCAGCTGGAGGCCACGTCGATTTTG 5223
 Db 67 CTGTGTACAGGCTGAGCCGCCAAAATTGGGGTTTCAGCTGGAGGCCACGTCGATTTTG 8
 QY 5224 GCTTTGT 5230
 Db 7 GCTTTGT 1

RESULT 15
 ABV97891/C
 ID ABV97891 standard; cDNA; 416 BP.
 AC ABV97891;
 XX
 DT 14-JAN-2003 (first entry)
 XX
 DE Human pancreatic cancer expressed cDNA SEQ ID NO 3299.
 KW Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;
 KW cytostatic; tumour; gene, ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200260317-A2.
 XX
 PD 08-AUG-2002.
 XX
 PF 30-JAN-2002; 2002WO-US02781.
 XX
 PR 30-JAN-2001; 2001US-265305P.
 PR 31-JAN-2001; 2001US-265682P.
 PR 09-FEB-2001; 2001US-267588P.
 PR 21-MAR-2001; 2001US-278651P.
 PR 28-APR-2001; 2001US-287112P.
 PR 16-MAY-2001; 2001US-291631P.
 PR 12-JUL-2001; 2001US-305484P.
 PR 20-AUG-2001; 2001US-313999P.
 PR 27-NOV-2001; 2001US-333626P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Benson DR, Kalos MD, Lodes MJ, Persing DH, Hepler WT, Jiang Y;
 DR WPI; 2002-627435/67.
 XX
 PT New isolated polynucleotide and pancreatic tumor polypeptides, useful
 PT for diagnosing, preventing and/or treating cancer, particularly
 PT pancreatic cancer.
 XX
 PS Claim 1; SEQ ID NO 3299; 300pp + Sequence Listing; English.
 XX

CC The invention relates to an isolated polynucleotide (I) comprising: (a)
 CC any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145);
 CC (b) complements of (a); (c) sequences consisting of at least 20
 CC contiguous residues of (a); (d) sequences that hybridize to (a), under
 CC moderately stringent conditions; (e) sequences having at least 75% or 90%
 CC identity to (a); or (f) degenerate variants of (a). Polypeptides
 CC (ABP68596-ABP68637) encoded by (I) and oligonucleotide can be used to
 CC detect cancer in a patient and compositions comprising polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations and
 CC antigen presenting cells expressing the polypeptide are useful in
 CC treating pancreatic cancer and stimulating an immune response. The
 CC polynucleotides can be used as probes or primers for nucleic acid
 CC hybridisation, in the design and preparation of ribozyme molecules for
 CC inhibiting expression of the tumour polypeptides and proteins in the
 CC tumour cells, in vaccines and for gene therapy.
 CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 416 BP; 101 A; 113 C; 114 G; 88 T; 0 other;
 Query Match 7.6%; Score 414.4; DB 24; Length 416;
 Best Local Similarity 99.8%; Pred. No. 1.1e-92;
 Matches 415; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4815 TATCAACACGAGCCCTTGACCCCTTGAGTCTGGCTTTGGCTGACCCCTCTTTGGGTCTCA 4874
 Db 416 TATCAACACGAGCCCTTGACCCCTTGAGTCTGGCTTTGGCTGACCCCTCTTTGGGTCTCA 357
 QY 4875 GTTCTCTCTCTGCAAAACAAGTTGCCATCTGTTTGGCTTCCAGCACATAAAGTAATGAA 4934
 Db 356 GTTCTCTCTCTGCAAAACAAGTTGCCATCTGTTTGGCTTCCAGCACATAAAGTAATGAA 297
 QY 4935 CTTTGATGATGCTTTTGTCTGGGCAATTATGTGTCCATGCCAGGGATGCCACAGGGGCCCC 4994
 Db 296 CTTTGATGATGCTTTTGTCTGGGCAATTATGTGTCCATGCCAGGGATGCCACAGGGGCCCC 237
 QY 4995 AGTCCAGGTGGCTTAACAGCATCTCAGGGAATGTCATCTGGAGCTGGCAAGACCCCTGC 5054
 Db 236 AGTCCAGGTGGCTTAACAGCATCTCAGGGAATGTCATCTGGAGCTGGCAAGACCCCTGC 177
 QY 5055 AGACCTCATAGAGCCTCATCTGGTGGCCACAGCAGCCAAAGCTAGAGCCCTCCGGATCCC 5114
 Db 176 AGACCTCATAGAGCCTCATCTGGTGGCCACAGCAGCCAAAGCTAGAGCCCTCCGGATCCC 117
 QY 5115 ATCCAGGCGCAAGAGGAATAGGAGGACATAGGAGGACATGGAAACCATTTTGCCTCTGGCTGTGTACAG 5174
 Db 116 ATCCAGGCGCAAGAGGAATAGGAGGACATAGGAGGACATGGAAACCATTTTGCCTCTGGCTGTGTACAG 57
 QY 5175 GGTGAGCCCCAAAATTGGGGTTTCAGCTGGAGGCCACGTCGATTTCTTTGGCTTTGT 5230
 Db 56 GGTGAGCCCCAAAATTGGGGTTTCAGCTGGAGGCCACGTCGATTTCTTTGGCTTTGT 1
 Search completed: January 29, 2004, 03:29:21
 Job time : 1357.51 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2004, 03:01:52 ; Search time 10849.3 Seconds
(without alignments)
12195.625 Million cell updates/sec

Title: US-09-996-617-1
Perfect score: 5444
Sequence: 1 gccccaggccctggagaggt.....aggaataagaagttacctac 5444

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pbg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	994	18.3	1201	9	AL541851
c	2	983.2	18.1	1201	13 BX341134
3	898.4	16.5	1201	13	BX341135
4	819.8	15.1	1055	13	BQ072636

ALIGNMENTS

RESULT 1
AL541851

LOCUS

DEFINITION

AL541851 Homo sapiens PLACENTA

5-PRIME, mRNA sequence.

ACCESSION

AL541851

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AL541851 1201 bp mRNA linear EST 12-MAY-2003
AL541851 Homo sapiens PLACENTA Homo sapiens CDNA clone CS0DE007D22
5-PRIME, mRNA sequence.
AL541851
AL541851.2 GI:30546421
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12873320.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 8396.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DE007D22
Feng Liang Email : fliang@lifetech.com URL :

5	814.4	15.0	1046	12	BM549709
6	787.4	14.5	1110	13	EX457482
7	767	14.1	768	12	BI518887
8	749.4	13.8	894	14	CD515485
9	727	13.4	791	10	BG682080
c	10	722.4	764	13	BQ576345
11	671	12.3	686	14	CA308213
12	633.6	11.6	672	12	BI911134
13	633	11.5	728	12	BI908442
c	14	628.4	655	12	BQ003173
15	623.6	11.5	655	12	BM152462
16	615	11.3	615	12	BM791087
c	17	575	592	13	BQ633947
c	18	566	566	10	BF059271
19	558.4	10.3	560	12	BM152839
20	544.4	10.0	554	12	BM193542
c	21	520.8	524	9	AW173055
22	520.4	9.6	523	12	BM153182
23	509	9.3	595	13	BQ355077
24	502.6	9.2	514	9	AV682872
25	495.4	9.1	580	12	BM152807
c	26	494	494	9	AW593727
c	27	489.6	806	12	BI519523
c	28	485	494	12	BM151763
c	29	483.4	500	10	BE775300
30	476.4	8.8	633	12	BI909022
31	475	8.7	509	12	BG774227
c	32	474.6	521	9	AI819334
c	33	469.8	504	9	AI742596
c	34	469	469	9	AI825353
c	35	458	458	10	BE464469
c	36	456.8	511	13	BX474531
c	37	449	449	9	AI365110
c	38	445.4	461	9	AW291804
c	39	445.2	487	12	BI028123
c	40	441	464	9	AW291769
c	41	438	441	9	AI339112
c	42	434.4	466	12	BM193358
c	43	434.2	800	791	28 B2129303
c	44	424	425	9	AI439770
45	423.2	7.8	489	12	BM151640

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DE007DB11Q1.

FEATURES

source	1. .1201	/organism="Homo sapiens"	/mol_type="mRNA"	/db_xref="taxon:9606"	/clone="CS0DE007YD22"	/tissue_type="PLACENTA"	/clone_lib="Homo sapiens PLACENTA"	/notes="Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."	BASE COUNT	264 a	370 c	301 g	225 t	41 others	ORIGIN
		Query Match	18.3%	Score 994;	DB 9;	Length 1201;									
		Best Local Similarity	96.2%	Pred. No. 9.9e-124;											
		Matches 1008; Conservative	17;	Mismatches 22;	Indels 1;	Gaps 1;									
Qy	5	CAGGGCC	TGAGAGGT	CTGAAGAA	ACTGGAG	CAGCAGC	CCGGGGCTCCACT	CTGGGT	64						
Db	32	CCGGT	CCGGAAT	TCCCGGAT	GAAAC	CTGGAG	CCAGCAGC	CCGGGGCTCCACT	GGGT	91					
Qy	65	TCTGAA	AGCCATT	TCCCTGCT	CTGCGCT	CCTCC	CACCCACCT	CTTCT	CAGCCT	TGCAG	124				
Db	92	TCTGAA	AGCCATT	TCCCTGCT	CTGCGCT	CCTCC	CACCCACCT	CTTCT	CAGCCT	TGCAG	151				
Qy	125	CTCAAG	GGTTGAT	TCTCAGAG	GTCCAG	ACCAGG	AGGAAAGAA	ATCTG	AGGAAC	ACACAGA	184				
Db	152	CTCAAG	GGTTGAT	CTCAGG	AGTCCAG	ACCAGG	AGGAAAGAA	ATCTG	AGGAAC	ACACAGA	211				
Qy	185	ACAGT	AGCGTT	TGCCAC	ACCCAT	CTCCCGT	CACACAT	CTCCCT	CACCCCT	CACCCCTC	244				
Db	212	ACAGT	AGCGTT	TGCCAC	ACCCAT	CTCCCGT	CACACAT	CTCCCT	CACCCCT	CACCCCTC	271				
Qy	245	CTGCGC	TGGCCCT	TGGAC	CCCCAT	CCAGG	ACCTCCCT	TATCAG	CTGACT	TTCTTC	CAGTGT	304			
Db	272	CCTG	CCCTGG	CCCTGG	ACCCCAT	CCAGG	ACCTCCCT	TATCAG	CTGACT	TTCTTC	CAGTGT	331			
Qy	305	TGACG	CCCCCT	CTGGGT	CTCTCC	CTCCCT	TGGCTTTT	CTAC	CACCTCC	CCCTCT	TATCGC	364			
Db	332	TGACG	CCCCCT	CTGGGT	CTCTCC	CTCCCT	TGGCTTTT	CTAC	CACCTCC	CCCTCT	TATCGC	391			
Qy	365	GTCT	ATCTGT	GTAGTG	CGCTCG	GGATTT	TATAAA	CTGGG	TCCGAAT	CTGTAAT	TAAGAG	424			
Db	392	GTCT	ATCTGT	GTAGTG	CGCTCG	GGATTT	TATAAA	CTGGG	TCCGAAT	CTGTAAT	TAAGAG	451			
Qy	425	GTAAG	AGCCAA	GCAAG	CAG	CACTGT	TCTCTG	CTCG	CTGAT	ACCTC	CACCA	CTGG	484		
Db	452	GTAAG	AGCCAA	GCAAG	CAG	CACTGT	TCTCTG	CTCG	CTGAT	ACCTC	CACCA	CTGG	511		
Qy	485	GAAC	ATCCCC	CCAG	ACCCCT	CTTAA	CTCCGG	GACAG	ATGGCT	TGGCGG	AGCCTGG	GGCC	544		
Db	512	GAAC	ATCCCC	CCAG	ACCCCT	CTTAA	CTCCGG	GACAG	ATGGCT	TGGCGG	AGCCTGG	GGCC	570		
Qy	545	GCCT	TGGCT	GTACT	TGGAGT	TTCTG	TGAAG	AGGAGG	AGCTG	AAAGAGT	TTCCAG	CTTCTG	604		
Db	571	GCCT	TGGCT	GTACT	TGGAGT	TTCTG	TGAAG	AGGAGG	AGCTG	AAAGAGT	TTCCAG	CTTCTG	630		
Qy	605	TCGCG	CAATAA	AGCG	CACTCC	CAGG	AGCTCTTT	CGGGT	TGAG	ACCCG	CTCAG	CCAG	AGA	664	
Db	631	TCGCG	CAATAA	AGCG	CACTCC	CAGG	AGCTCTTT	CGGGT	TGAG	ACCCG	CTCAG	CCAG	AGA	690	
Qy	665	CGAGT	CGCAT	TGAG	GTGG	CGCTCG	TACT	TGGG	AGCAG	CGCG	CGCTGG	724			
Db	691	CGAGT	CGCAT	TGAG	GTGG	CGCTCG	TACT	TGGG	AGCAG	CGCG	CGCTGG	750			
Qy	725	ACCT	AGCCCT	TCCATA	CTTGG	AGCAG	ATGGG	GTGAG	GTCACT	GTG	CGCC	CAAG	CC	AGG	784
Db	751	ACCT	AGCCCT	TCCATA	CTTGG	AGCAG	ATGGG	GTGAG	GTCACT	GTG	CGCC	CAAG	CC	AGG	810

Qy	785	AAGGGGAGGGC	CATCTCTCCCTCATTTCCCTCTACAGCCCAAGTGAACCCACCTGGGGTCTC	844
Db	811	AAGGGGAGGGC	CACCTCTCTCCCTCATTTCCCTCTACAGCCCAAGTGAACCCACCTGGGGTCTC	870
Qy	845	CCAGCCAAACCA	CACTCCACCCGACGTGCTAATGCCCTGGATCCATGAATTGCCGGCGGGT	904
Db	871	CCAGCCAAACCA	CTCCACCCGACGTGCTAATGCCCTGGATCCATGAATTGCCGGCGGGT	930
Qy	905	GCACCCAGGGCT	TCAGAGAGAAAGGGTTTTTGAGACAGCTGCCTGACACATCTGGACGCCGT	964
Db	931	GCACCCAGGGCT	TCAGAGARAAGGGTTTTTGAGAMARCTGCCTGACACATCTGGACGCGST	990
Qy	965	GGAGAGAAATCT	CTGCGCTCACTCTCTACCAAGCTCTTCCAAGCTCCCAGACCATGAGT	1021
Db	991	GGAGAGAAATCT	TTKCTTACTTCTTACCAAGCTCTTCAAGCTCCCAACCATGAST	1051
Qy	1025	CTTCAAGCCAGGAGT	CACCCAAAGCCCC	1052
Db	1051	TTTCAACAGGGRKT	TACCAMGCCCCCC	1078

RESULT 2
BX341134/c

EX3411134/C	EX3411134	1201 bp	linear	EST 01-MAY-2003
LOCUS	EX3411134	1201 bp	linear	EST 01-MAY-2003
DEFINITION	EX3411134 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED	1201 bp	linear	EST 01-MAY-2003
	Homo sapiens cDNA clone CS0DJ012YC02 3-PRIME, mRNA sequence.			
ACCESSION	EX3411134			
VERSION	EX3411134.1	GI:30309318		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
	Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 1201)			
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.			
TITLE	Full-length cDNA libraries and normalization			
JOURNAL	Unpublished			
COMMENT	Contact: Genoscope			

COMMENT

Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 131 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 8396.r For
more information about this cluster, see
<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DJ012BB01NP1>
sccluster=8396.r. Contact : Feng Liang Email : liang@lifetech.com
URL : <http://fulllength.invitrogen.com/Invitrogen/Corporation1600>
Paradise Avenue Genoscope sequence ID : CS0DJ012BB01NP1.

FEATURES

```

FEATURES
source
1. .i201
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODJ012YC02"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
263 286 288 313 51 others
BASE COUNT

```

BASE CO
ORIGIN

Query Match	18.1%	Score 983.2;	DB 13;	Length 1201;
Best Local Similarity	95.6%;	Pred. No. 2.8e-122;		
Matches 1000;	Conservative	9;	Mismatches 37;	Indels 0;
				Gaps 0;

Qy	4366	GCCCACTTGGGATCAGGATCAGGCTGCAAGTGAAGACAA	GAAAGATGAGACTCTGGTG	4425
Db	1057	GTCYACKTGCCACTTGGATCRGGATCAGCTGCAAGTGAAR	CAAGAAAGWGAGACTCTGG	998

QY	4426	TGGGAGGCTTGGTGAACACAGAGAGATCTATGCTGCAACTACTCTGATCCCTCCAGCC	4485
Db	997	TGTGAGSCTTGTGAAACCARAGATCTCATGCTGCAACTACTCTGATCCCTCCAGCC	938
QY	4486	CGCATAGCGTACCTTACCTCTGATGCCCCGCGAGTTGCTGCTACTTTGTGACAGTAT	4545
Db	937	CGCATAGCGTACCTTACCTCTGATGCCCCGCGAGTTGCTGCTACTTTGTGACAGTAT	878
QY	4546	CGAGAGCAGCTATAGCCGAGTGCATCGGTGGAGGTGTCTTGGACAAACTGATGGA	4605
Db	877	CGAGAGCAGCTATAGCCGAGTGCATCGGTGGAGGTGTCTTGGACAAACTGATGGA	818
QY	4606	CAGGTGCTGAGCCAGGAGCAGTACGAGAGGTGCTGGCTGAGAAACAGAGGCCAGCAG	4665
Db	817	CAGGTGCTGAGCCAGGAGCAGTACGAGAGGTGCTGGCTGAGAAACAGAGGCCAGCAG	758
QY	4666	ATGCGGAAGCTTTGAGTGGAGCCAGTCTGCGGACCGGAGTGCAGAGTGCATCTAC	4725
Db	757	ATGCGGAAGCTTTGAGTGGAGCCAGTCTGCGGACCGGAGTGCAGAGTGCATCTAC	698
QY	4726	CAAGCCCTGAAGAGACCCATCTCACTCTATTATGGAATCTGCGAGAGGGCAGCAAA	4785
Db	697	CAAGCCCTGAAGAGACCCATCTCACTCTATTATGGAATCTGCGAGAGGGCAGCAAA	638
QY	4786	AAGGGAATCTCTGCACTCAGCAGCTGAAGTATCAACACAGCCCTTGACCCCTTGATCT	4845
Db	637	AAGGGAATCTCTGCACTCAGCAGCTGAAGTATCAACACAGCCCTTGACCCCTTGATCT	578
QY	4846	GGCTTTGGCTGACCTCTTTTGGTCTGAGTTCTTTCTCTGCAACAGTTGCCATCTG	4905
Db	577	GGCTTTGGCTGACCTCTTTTGGTCTGAGTTCTTTCTCTGCAACAGTTGCCATCTG	518
QY	4906	GTTCGCTTCCAGCACTAAAGTAATGGAATTTGATGATGCTTTGCTGGCATTTATG	4965
Db	517	GTTCGCTTCCAGCACTAAAGTAATGGAATTTGATGATGCTTTGCTGGCATTTATG	458
QY	4966	TCCATGCCAGGATGCAACAGGGGGCCCGAGTCCAGGTGGCTTAACAGCATCTCAGGAA	5025
Db	457	TCCATGCCAGGATGCAACAGGGGGCCCGAGTCCAGGTGGCTTAACAGCATCTCAGGAA	398
QY	5026	TGTCCATCTGGAGCTGGGAAGACCCCTGCGAGCTCATAGAGCTCATCTGTTGGCCACA	5085
Db	397	TGTCCATCTGGAGCTGGGAAGACCCCTGCGAGCTCATAGAGCTCATCTGTTGGCCACA	338
QY	5086	GCAGCCAGCCTAGAGCCCTCCGATCCCATCCAGGCGCAAGAGAAATAGAGGACAT	5145
Db	337	GCAGCCAGCCTAGAGCCCTCCGATCCCATCCAGGCGCAAGAGAAATAGAGGACAT	278
QY	5146	GGAAACATTTGCTCTGGCTGTGTCAACAGGGTGAGCCCAAAATTTGGGTTTTCAGCTGG	5205
Db	277	GGAAACATTTGCTCTGGCTGTGTCAACAGGGTGAGCCCAAAATTTGGGTTTTCAGCTGG	218
QY	5206	AGGCCAGCTGGATTTTGGCTTTGTACAGAGATCTCAAGAGCAAGCCCAACAGATAA	5265
Db	217	AGGCCAGCTGGATTTTGGCTTTGTACAGAGATCTCAAGAGCAAGCCCAACAGATAA	158
QY	5266	ACTGGAAGGAATTTTATTCAGAAATAAAGGAGTATCAGAGCTCTTTTGAATTTGCTTA	5325
Db	157	ACTGGAAGGAATTTTATTCAGAAATAAAGGAGTATCAGAGCTCTTTTGAATTTGCTTA	98
QY	5326	GCAGGCTTTTCCAGTTTATTCAGAAACCCCTTATAAATTTTAAATTTTCTTAAATTT	5385
Db	97	GCAGGCTTTTCCAGTTTATTCAGAAACCCCTTATAAATTTTAAATTTTCTTAAATTT	38
QY	5386	AAGAATTAATAAATAACAAAAAGAA	5411
Db	37	AAGAATTAATAAATAACAAAAAGAA	12

RESULT 3
BX341135
LOCUS

BX341135 1201 bp mRNA linear EST 01-MAY-2003

DEFINITION	BX341135 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED	
ACCESSION	BX341135	
VERSION	BX341135.1 GI:30310958	
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS	1. (bases 1 to 1201)	
TITLE	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.	
JOURNAL	Full-length cDNA libraries and normalization	
COMMENT	Unpublished Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 8396.r For more information about this cluster, see http://www.genoscope.cns.fr/ cli-bin/cluster.cgi?seq=CS0DJ012BB01QP1&cluster=8396.r. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DJ012BB01QP1.	
FEATURES	Location/Qualifiers 1..1201 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0DJ012YC02" /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED" /cell_line="JURKAT" /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED" /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."	
BASE COUNT	297 a	329 c 312 g 217 t 46 others
ORIGIN		
Query Match	16.5%;	Score 898.4; DB 13; Length 1201;
Best Local Similarity	97.4%;	Pred. No. 6.2e-111;
Matches	937; Conservative	7; Mismatches 13; Indels 5; Gaps 3;
QY	773	CCCAGCCCGAGAGGGGCGAGCCACTCTCCCTCAATCCCTACAGCCCAAGTGAACCCC 832
Db	57	CCGGGATCAGAGAGGGGCGAGCCACTCTCCCTCAATCCCTACAGCCCAAGTGAACCCC 116
QY	833	ACCTGGGGTCTCCAGCAACCCACCTCCAGCAGTGTATGCTCCCTGGATCCATGAAT 892
Db	117	ACCTGGGGTCTCCAGCAACCCACCTCCAGCAGTGTATGCTCCCTGGATCCATGAAT 176
QY	893	TGCCGGGGGTGTACCCAGGGCTCAGAGAGAGGGTTTGTAGACAGTGCCTGCACACAT 952
Db	177	TGCCGGGGGTGTACCCAGGGCTCAGAGAGAGGGTTTGTAGACAGTGCCTGCACACAT 236
QY	953	CTGGAGCCGCTGGAGAGAAATCTCTGCTCACTCTCTACCAAGCTTCTCAAGTCCC 101
Db	237	CTGGAGCCGCTGGAGAGAAATCTCTGCTCACTCTCTACCAAGCTTCTCAAGTCCC 296
QY	1013	CAGACCATGAGTCTCCAGCCAGGAGTACCAACGCCGCCACATCCACAGCAGTCTGG 107
Db	297	CAGACCATGAGTCTCCAGCCAGGAGTACCAACGCCGCCACATCCACAGCAGTCTGG 356
QY	1073	GGAGCTGGGGATCCCACTCAGCCAGCCTAGCAGCCAGCAGAGAGGAGTCTCTGGGA 113
Db	357	GGAGCTGGGGATCCCACTCAGCCAGCCTAGCAGCCAGCAGAGAGGAGTCTCTGGGA 416
QY	1133	CCCAATGGCCTCTGGATGAACCGTCAGGAATTTACTACAGAAATCAGAGAAAGAGA 119
Db	417	CCCAATGGCCTCTGGATGAACCGTCAGGAATTTACTACAGAAATCAGAGAAAGAGA 476


```

QY 780 CCAGGAAGGGGAGGCGCACTCTCCCTCATTCCTCCCTACAGCCCAAGTGAACCCCACT-GG 838
Db 830 CAGGGAAGGGGAGGCGCACTCTCTCTCATTCCTCCCTAAAGCCCAAGTGACCCCACTGGG 889
QY 839 GGTCTCCCAAGCAACCC 855
Db 890 GGTCTCCCAAGCAACCC 906

RESULT 5
BM549709 1046 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT_6544160 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5745766
DEFINITION 5', mRNA sequence.
ACCESSION BM549709
VERSION BM549709.1 GI:18785339
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1046)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: Agencourt Bioscience Corporation
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12769 row: d column: 23
High quality sequence start: 14
High quality sequence stop: 649.
Location/Qualifiers
1. .1046
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5745766"
/tissue_type="leukocyte"
/lab_host="DH10B"
/clone_lib="NIH MGC 118"
/note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV
(destroyed); RNA source leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dr primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 027. Note:
this is a NIH MGC Library."
BASE COUNT 227 a 293 c 295 g 231 t
ORIGIN

Query Match 15.0%; Score 814.4; DB 12; Length 1046;
Best Local Similarity 98.4%; Pred. No. 1.1e-99;
Matches 844; Conservative 0; Mismatches 11; Indels 3; Gaps 2;

QY 3712 TCTCAAGGGGAGCTGCATAGAGCTTTGGGACCTGACGATCATTCTGGGGCCCCAGC 3771
Db 20 TCTCAAGGGGAGCTGCATAGAGCTTTGGGACCTGACGATCATTCTGGGGCCCCAGC 79
QY 3772 GGGCCCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTGACCGAGTTCACTTCCT 3831
Db 80 GGGCCCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTGACCGAGTTCACTTCCT 139
QY 3832 GTAGCTGGCTCTCAACCGTGGCCCAACACGGGTCTCTGCTTTGTGTATGAGAGAGCGGTG 3891
Db 140 GTAGCTGGCTCTCAACCGTGGCCCAACACGGGTCTCTGCTTTGTGTATGAGAGAGCGGTG 199

```

FEATURES

```

source
3892 ACCGTTGAGATTGAATTTCTGTGTGTGGACCAAGTTCTTGGGTGAGATCAACCCACAGCAC 3951
200 ACCGTTGAGATTGAATTTCTGTGTGTGGACCAAGTTCTTGGGTGAGATCAACCCACAGCAC 259
3952 AGCTGGATGGTGGCAGGGCCCTCTCTGACATCAAGCTGAGCTGAGCTGGAGCTGGAAGCT 4011
260 AGCTGGATGGTGGCAGGGCCCTCTCTGACATCAAGCTGAGCTGAGCTGGAGCTGGAAGCT 319
4012 GTGACCTCCCTCACTTTGTGGCTCTCCAAAGGGGCGCATGTGGACACATCCCTGTTCCAA 4071
320 GTGACCTCCCTCACTTTGTGGCTCTCCAAAGGGGCGCATGTGGACACATCCCTGTTCCAA 379
4072 ATGGCCCACTTTAAAGAGAGGGGATGTCTCTGGAGAGCCAGCCAGGGTGAGCTGCAT 4131
380 ATGGCCCACTTTAAAGAGAGGGGATGTCTCTGGAGAGCCAGCCAGGGTGAGCTGCAT 439
4132 CACATAGTTCTGGAAAAACCCAGCTTCTCCCTTGGGAGTCTCTGAAATGATCCAT 4191
440 CACATAGTTCTGGAAAAACCCAGCTTCTCCCTTGGGAGTCTCTGAAATGATCCAT 499
4192 AATGCCCTCGGCTTCATTTCCCGTCACTCTGTGGTGTGCTTTACACCGCTCCATCT 4251
500 AATGCCCTCGGCTTCATTTCCCGTCACTCTGTGGTGTGCTTTACACCGCTCCATCT 559
4252 GAGGAAGTCACTTCCACCTCTTACCTGATCCCAAGTGAAGTCTCTCCATTCGGAAGAACTG 4311
560 GAGGAAGTCACTTCCACCTCTTACCTGATCCCAAGTGAAGTCTCTCCATTCGGAAGAACTG 619
4312 GAGCTCTGTATCGAAGCCCTGGAGAACACAGCTGTTCTCGGAGTTCTACGTTGGCCAC 4371
620 GAGCTCTGTATCGAAGCCCTGGAGAACACAGCTGTTCTCGGAGTTCTACGTTGGCCAC 679
4372 TTGGGATCAGGATCAGGCTGCAAGTGAAGAACAGAGATGAGACTCTGCT-GTGGGA 4430
680 TTGGGATCAGGATCAGGCTGCAAGTGAAGAACAGAGATGAGACTCTGCTGTGGTGGGA 739
4431 GGCCTTGGTGAACCCAGGAGATCTCATGCTGCAACTACTCTGATCCCTCCAGCCCGCAT 4490
740 GGCCTTGGTGAACCCAGGAGATCTCATGCTGCAACTACTCTGATCCCTCCAGCCCGCAT 799
4491 ACCGTTACTTCACTCTGGATGCCCGGAGTTGC--TGCACCTTTGTGGACAGTATCGA 4548
800 AGGCTACTCTTCACTCTGGATGCCCGGAGTTGCTGCTTGGGGAGCAGTATCGA 859
4549 GAGCAGCTGATAGCCCGA 4566
860 GAGCAGCTGATAGCCCGA 877

RESULT 6
EX457482 1110 bp mRNA linear EST 22-MAY-2003
LOCUS EX457482 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP008YN15
DEFINITION 5-PRIME, mRNA sequence.
ACCESSION EX457482
VERSION EX457482
KEYWORDS EX457482.1 GI:31028889
SOURCE EST.
ORGANISM Homo sapiens (human)
REFERENCE 1 (bases 1 to 1110)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 8396.r For
more information about this cluster, see

```


http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0CAP008CG08QPl&cluster=8396.r. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/InvitroGen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0CAP008CG08QPl.

FEATURES

source
1. .1110
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0CAP008YNI5"
/issue_type="THYMUS"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
Library was not normalized."
BASE COUNT 218 a 270 c 257 g 262 t 103 others
ORIGIN
Query Match 14.5%; Score 787.4; DB 13; Length 1110;
Best Local Similarity 88.3%; Pred. No. 4.6e-96;
Matches 848; Conservative 58; Mismatches 48; Indels 6; Gaps 6;
QY 1644 GTCCAAAGGTGGTGAAGTCTCGTCTGAGTCTCATCGGAAAGATGGACAGCCACTCCGGCTCC 1703
DB 72 GACCCACGGTCGGGTCCTCGTCTGAGTCTCATCGGAAAGATGGACAGCCACTCCGGCTCC 131
QY 1704 CATTAGACAGATCTCTAGGCGAGCGGGCTCTCTTCATCTCTCGATGGTGTAGATGA 1763
DB 132 CATTAGACAGATCTCTAGGCGAGCGGGCTCTCTTCATCTCTCGATGGTGTAGATGA 191
QY 1764 GCAGATGGGTCTTCAGAGCGCGAGTCTGAGCTCTGCTGCACTGGAGCCAGCCACA 1823
DB 192 GCAGATGGGTCTTCAGAGCGCGAGTCTGAGCTCTGCTGCACTGGAGCCAGCCACA 251
QY 1824 GCGCGCGGATGCACTCTGGGAGTTTGTCTGGGAAATCTATCTTCCCGAGGCACTCCCT 1883
DB 252 GCGCGCGGATGCACTCTGGGAGTTTGTCTGGGAAATCTATCTTCCCGAGGCACTCCCT 311
QY 1884 CTTGATCAGCGCTCGGACACAGCTCTGAGAACCTTCATCTCTCTTTGGAGCAGCAGC 1943
DB 312 CTTGATCAGCGCTCGGACACAGCTCTGAGAACCTTCATCTCTCTTTGGAGCAGCAGC 371
QY 1944 TTGGGTAGAGTCTCTGGGTTCTCTGAGTCCAGCAGGAAGTAATTTCTACAGATATTT 2003
DB 372 TTGGGTAGAGTCTCTGGGTTCTCTGAGTCCAGCAGGAAGTAATTTCTACAGATATTT 431
QY 2004 CACAGATGAAGGCAAGCAATTAGAGCTTTAGTTGGTCAATCAAAACAAAGAGCTCTG 2063
DB 432 CACAGATGAAGGCAAGCAATTAGAGCTTTAGTTGGTCAATCAAAACAAAGAGCTCTG 491
QY 2064 GGCCCTGTGTCTTGTGCTCCCTGGGTGCTCTGGTGGCTGCACTGTGCTGATGAGCAGAT 2123
DB 492 GGCCCTGTGTCTTGTGCTCCCTGGGTGCTCTGGTGGCTGCACTGTGCTGATGAGCAGAT 551
QY 2124 GAAGCGGAAGGAAAACTCAGCTGACTTCCAGACCAACACACCTCTGTCTACATTA 2183
DB 552 GAAGCGGAAGGAAAACTCAGCTGACTTCCAGACCAACACACCTCTGTCTACATTA 611
QY 2184 CTTTGGCCAGGCTCTCAAGCTCAGCCATTGGGACCCAGCTCAGAGA-CCTCTGCTCTC 2242
DB 612 CTTTGGCCAGGCTCTCAAGCTCAGCCATTGGGACCCAGCTCAGAGA-CCTCTGCTCTC 671
QY 2243 TGGCTCTCAGGGCATCTGGCAAAA-AAGACCTTTTTCAGTCCAGATGACCTCAGGAAG 2301
DB 672 TGGCTCTCAGGGCATCTGGCAAAA-AAGACCTTTTTCAGTCCAGATGACCTCAGGAAG 731
QY 2302 CATGGG-TTAGATGGGGCCATCATCTCCAGCTTTTGAAGATGGGTATTTCTTCAAGAGCA 2360
DB 732 CATGGGTTAGATGGGGSCATCATCTCCACCTTTTGAAGATGGGTATTTCTTCAAGAGCA 791

QY 2361 CCCATCCCTCTGAGCTACAGCTTCATTCACTCTGTTTCCAAAGATT-CTTTGAGCAA 2419
DB 792 CCCATCCCTCTGAGCTACAGCTTCATTCACTCTGTTTCCAAAGATT-CTTTGAGCAA 851
QY 2420 TGTCTATCTCTGAGGATGAGAAAGGAGGAGGTAACATCTTAATTCATCATAGA-T 2478
DB 852 TGTCTATCTCTGAGGATGAGAAAGGAGGAGGTAACATCTTAATTCATCATAGA-T 911
QY 2479 TTGGAAGAAGACGCTAGA-AGCATATGAGAAATCATGCGCTGTTTGGGCGATCAACACACG 2537
DB 912 TTGGAAGACGCTAGA-AGCATATGAGAAATCATGCGCTGTTTGGGCGATCAACACACG 971
QY 2538 TTTCTATTTGGGCTCTTAAAGTGTAGTGGGAGAGAGAGATGGAGACATCTTTCACTG 2597
DB 972 TTTGATMCWGGGCTTTTTCVTCWCKASKGWCACMCMTKTCCTMCACCTTMCAMCKTTG 1031
RESULT 7
BI518887 768 bp mRNA linear EST 29-AUG-2001
LOCUS 603061807F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5211072 5',
DEFINITION mRNA sequence.
ACCESSION BI518887
VERSION BI518887.1 GI:15343679
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 768)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9pbbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Prepared by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11529 row: n column: 01
High quality sequence stop: 766.
FEATURES
Location/Qualifiers
1. .768
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5211072"
/tissue_type="leukocyte"
/lab_host="DH10B"
/clone_lib="NIH_MGC_118"
/note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH MGC Library."
BASE COUNT 176 a 213 c 225 g 154 t
ORIGIN

Query Match 14.1%; Score 767; DB 12; Length 768;
Best Local Similarity 100.0%; Pred. No. 2.7e-93;
Matches 767; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3119 CCCTGACGAGCTGAGCTGAGCTTCATCTGCTCAGGATGCTGAGCAGCAACACCTTT 3178
DB 1 CCCTGACGAGCTGAGCTGAGCTTCATCTGCTCAGGATGCTGAGCAGCAACACCTTT 60


```

|||||
661 GGTCTCTGCTTTGTGATGAGAGAGCGGTGACCGTTGAGATTGAATTCGTGTGGGAC 720
|||||
3922 CAGTT--CTGGGTGAGATCAACCCACAGCAGCTGAGTGTGGCAGGCGCTCTGCTGG 3979
|||||
721 CAGTTTCTGGGGTGAAGTCAACCCACAGCAGCAGCTGGGATGGTGCGAGGCGCCCTTGGC 780
|||||
3980 ACATCAAGGCTGAGCTGGAGCTGTGGAAGCTGTG 4014
|||||
781 TGGNACATCAAGGGCTTGAACCTGGGAAGCTGTG 815
|||||

RESULT 9
BG682080
LOCUS
DEFINITION
602630057F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4754607 5',
mRNA sequence.
ACCESSION
BG682080
VERSION
BG682080.1 GI:13913477
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 791)
NIH-MGC http://mgs.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10616 row: b column: 16
High quality sequence stop: 770.
FEATURES
Location/Qualifiers
1..791
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4754607"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Skn4"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 195 a 203 c 200 g 193 t
ORIGIN

Query Match 13.4%; Score 727; DB 10; Length 791;
Best Local Similarity 98.2%; Pred. No. 6.1e-88;
Matches 777; Conservative 0; Mismatches 10; Indels 4; Gaps 4;

QY 2045 AATCAACAAAGAGCTCTGGCGCTGTGCTGTGTGTCCTGGTGTGCTGGCTGTGCA 2104
DB 3 AATCACAAAGAGCTCTGGCGCTGTGCTGTGTGTCCTGGTGTGCTGGCTGTGCA 61
QY 2105 CTTGCTGTATGACGAGATGAAGCGAAGAAAATCACTAGCTTCCAAAGACACCA 2164
DB 62 CTTGCTGTATGACGAGATGAAGCGAAGAAAATCACTAGCTTCCAAAGACACCA 121
QY 2165 CACCCCTGTCTACATTACCTTGGCCAGGCTCTCAAGCTCAGCATTGGGACCCAGC 2224
DB 122 CACCCCTGTCTACATTACCTTGGCCAGGCTCTCAAGCTCAGCATTGGGACCCAGC 181
QY 2225 TCAGAGACCTCTGCTCTGCTGTGCTGTGAGGCGATCTGGCAAAAAAGACCCCTTTTCAGTC 2284
|||||

```

```

DB 182 TCAGAGACCTCTGCTCTGCTGCTGAGGGCATCTGCAAAAAAGACCCCTTTTCAGTC 241
QY 2285 CAGATGACCTCAGGAAGCATGGTTAGATGGGGCCATCATCTCCACCTTTCTTTGAAGATGG 2344
DB 242 CAGATGACCTCAGGAAGCATGGTTAGATGGGGCCATCATCTCCACCTTTCTTTGAAGATGG 301
QY 2345 GTATTCTTCAAGAGCACCCCATCCCTCTGAGCTACAGCTTCATTCACTCTGTTTCCAAAG 2404
DB 302 GTATTCTTCAAGAGCACCCCATCCCTCTGAGCTACAGCTTCATTCACTCTGTTTCCAAAG 361
QY 2405 AGTTCTTTTGACGAATGTCCTATGCTTTGGAGGATGAGAAGGGGAGAGGTAACATTTCTA 2464
DB 362 AGTTCTTTTGACGAATGTCCTATGCTTTGGAGGATGAGAAGGGGAGAGGTAACATTTCTA 421
QY 2465 ATTGCATCATAGATTGGAAGAGCGCTAGAGCATATGGAATACATCGCCCTGTTTGGGG 2524
DB 422 ATTGCATCATAGATTGGAAGAGCGCTAGAGCATATGGAATACATCGCCCTGTTTGGGG 481
QY 2525 CATCAACCACACGTTTCTTATTGGGCCCTGTTTAAGTGTAGTGGGGGAGAGAGATGGAGA 2584
DB 482 CATCAACCACACGTTTCTTATTGGGCCCTGTTTAAGTGTAGTGGGGGAGAGAGATGGAGA 541
QY 2585 ACATCTTTTCACTGCGGCTGTCTCAGGGGAGAAACCTGATGAGTGGGTCCCGTCCCTGC 2644
DB 542 ACATCTTTTCACTGCGGCTGTCTCAGGGGAGAAACCTGATGAGTGGGTCCCGTCCCTGC 601
QY 2645 AGCTGCTGCTCAGCCACACTCTCTGGAGTCCCTCCACTGCTGTACGAGACTCGGAACA 2704
DB 602 AGCTGCTGCTCAGCCACACTCTCTGGAGTCCCTCCACTGCTGTACGAGACTCGGAACA 661
QY 2705 AAAAGTTTCTTGACACAGTGTATGCCCATTTTCGAAGAAATGGGCA-TGTGTGTAGAAACA 2763
DB 662 AAA-GTTCTCTGACACAGTGTATGGGCCATTTTCGAAGAAATGGGCA-TGTGTGTAGAAACA 720
QY 2764 GACATGAGAGCTTTAGTGTGTCACCTTTCTGCAATT-AAATTCAGCCGCCACGTGAAGAAGCT 2822
DB 721 GACATGAGAGCTTTAGTGTGTCACCTTTCTGCAATTAAATTCAGCCGCCACGTGAAGAAGCT 780
QY 2823 TCAGCTGATTG 2833
DB 781 TCAGCTGATTG 791

RESULT 10
BG576345/c
LOCUS
DEFINITION
UI-H-E21-bbi-n-10-0-UI.s1 NCI_CGAP_Ch2 Homo sapiens cDNA clone
UI-H-E21-bbi-n-10-0-UI 3', mRNA sequence.
ACCESSION
BG576345
VERSION
BG576345.1 GI:21479662
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 764)
NCI_CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Dr. Steven Gitelis/ Rush Presbyterian, Dept. of
Orthopedics
JOURNAL
COMMENT
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bent-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 1-71, >POLY.#Simple repeat (matched complement) 72-100,
>AT richLow complexity 157-242, >MER41B#LTR/MER4-group
Seq primer: M13 FORWARD

```

Db	215	TACAAGACGACCCACAGAGTAAAGTGAAGGAGTATTTATTTCAGAAAAATAAAGGAGTAT	156	
Qy	5302	CACAGCTCTTTTAGAATTTGTCTAGCAGGCTTTCCAGTTTTCACAGAAAAACCCCTATAA	5361	
Db	155	CACAGCTCTTTTAGAATTTGTCTAGCAGGCTTTCCAGTTTTCACAGAAAAACCCCTATAA	96	
Qy	5362	ATTAATAATTTTCTTACTTAAATTTTAAAGATTTTAAAAAATACAAAAAGAAAAATGAAA	5421	
Db	95	ATTAATAATTTTCTTACTTAAATTTTAAAGATTTTAAAAAATACAAAAAGAAAAATGAAA	36	
Qy	5422	TAAAGGAATAAGAGTTA	5439	
Db	35	TAAAGGAATAAGAGTTA	18	
RESULT 11				
CA308213/c				
LOCUS				
DEFINITION				
UI-H-Ftl-bib-1-22-0-UI.s1 NCI CGAP_Ftl Homo sapiens cDNA clone				
CA308213				
ACCESSION				
VERSION				
CA308213.1 GI:24471267				
KEYWORDS				
EST.				
SOURCE				
Homo sapiens (human)				
ORGANISM				
Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE				
1 (bases 1 to 686)				
AUTHORS				
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.				
TITLE				
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),				
Tumor Gene Index				
JOURNAL				
COMMENT				
Unpublished				
Contact: Robert Strausberg, Ph.D.				
Email: cgapbs@mail.nih.gov				
Tissue Procurement: Dr. Gary W. Hunninghake, U of I				
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa				
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa				
Clone Distribution: Clone distribution information can be obtained				
from Dr. M. Bento Soares, bento-soares@uiowa.edu				
The following repetitive elements were found in this cDNA				
sequence: 32-117, >MER41B#LTR/MER4-group				
Seq primer: M13 FORWARD				
POLYA=Yes.				
FEATURES				
source				
1. 686				
/organism="Homo sapiens"				
/mol_type="mRNA"				
/db_xref="taxon:9606"				
/clone="UI-H-Ftl-bib-1-22-0-UI"				
/tissue_type="Aveolar Macrophage"				
/dev_stage="Adult"				
/lab_host="DH10B (Life Technologies)"				
/clone_lib="NCI CGAP_Ftl"				
/notes="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a				
modified polylinker; Site 1: Ecor I; Site 2: Not I;				
NCI CGAP_Ftl is a normalized cDNA library constructed from				
a pool of 81 RNA samples from Alveolar Macrophages				
challenged with different treatments. The library was				
normalized according to Bonaldo, Lennon and Soares, Genome				
Research, 6:791-806, 1996. First strand cDNA synthesis was				
primed with an oligo-dT primer containing a Not I site.				
Double stranded cDNA was ligated to an Ecor I adaptor,				
digested with Not I, and cloned directionally into				
pT73-Pac vector. The oligonucleotide used to prime the				
synthesis of first-strand cDNA contains a library tag				
sequence that is located between the Not I site and the				
(dT)18 tail. The sequence tag for this library is				
GGCCATGCGG. The tissue was provided by Dr. Gary W.				
Hunninghake of the University of Iowa.				
TAG LIB=UI-H-Ftl				
TAG TISSUE=Human Lung				
TAG_SEQ=GGCCATGCGG"				

POLYA=Yes.		Location/Qualifiers	
1. .764		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="UI-H-E21-bbi-n-10-0-UI"	
		/tissue_type="Chondrosarcoma Grade II"	
		/dev_stage="Adult"	
		/lab_host="DH10B (Life Technologies)"	
		/clone_lib="NCI CGAP Ch2"	
		/notes="Organ: Left Pelvis; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: Ecor I; Site 2: Not I; NCI CGAP_Ch2 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma Grade II. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is TGATCAGCT."	
		TAG LIB=UI-H-E21	
		TAG TISSUE=grade-2-chondrosarcoma	
		TAG_SEQ=ATCTAATATG"	
BASE COUNT	173 a	175 c	175 g
ORIGIN			
Query Match			
Best Local Similarity 99.5%; Pred. No. 2.5e-87;			
Matches 734; Conservative 0; Mismatches 3; Indels 1; Gaps 1;			
Qy	4702	CGAAGTGCAGATGACTTACCAAGCCCTGAAGGAGACCATCTCACTCATATG	4761
Db	754	CCGAAGTGCAGAGAT-GACTCTACCAAGCCCTGAAGGAGACCATCTCACTCATATG	696
Qy	4762	GAACTCTGGAGAGGCGCAAAAGGACTCTGCCACTCAGCAGCTGAAGTATCAAC	4821
Db	695	GAACTCTGNAGAAGGCGCAAAAGGACTCTGCCACTCAGCAGCTGAAGTATCAAC	636
Qy	4822	ACCAGCCCTTGACCTTGCTGCTTGGCTTGCTGACCCCTCTTTGGGCTCTCAGTTTCTT	4881
Db	635	ACCAGCCCTTGACCTTGCTGCTTGGCTTGCTGACCCCTCTTTGGGCTCTCAGTTTCTT	576
Qy	4882	TCTCTGAAACAAGTTGCCATCTGGTTTGGCTTCCAGCATTAAGTAATGGAACTTTGAT	4941
Db	575	TCTCTGAAACAAGTTGCCATCTGGTTTGGCTTCCAGCATTAAGTAATGGAACTTTGAT	516
Qy	4942	GATGCCCTTCTCGGCATATATGTCCATGCCAGGATGCCAGGGGCCCCCGTCCAG	5001
Db	515	GATGCCCTTCTCGGCATATATGTCCATGCCAGGATGCCAGGGGCCCCCGTCCAG	456
Qy	5002	GTGGCCTTAACAGCATCTCAGGGAATGCCATCTGGAGCTGGCAAGACCCCTGCAGACCTC	5061
Db	455	GTGGCCTTAACAGCATCTCAGGGAATGCCATCTGGAGCTGGCAAGACCCCTGCAGACCTC	396
Qy	5062	ATAGAGCCTCATCTGGTGGCCACAGCAGCCAGCCTAGAGCCCTCCGGATCCCATCCAG	5121
Db	395	ATAGAGCCTCATCTGGTGGCCACAGCAGCCAGCCTAGAGCCCTCCGGATCCCATCCAG	336
Qy	5122	CGAAAGAGAATAGAGGACATGGNACCATTTGCTCTGGCTGTGTACAGGGTACG	5181
Db	335	CGAAAGAGAATAGAGGACATGGNACCATTTGCTCTGGCTGTGTACAGGGTACG	276
Qy	5182	CCCAAAATTCGGGTTAGCGTGGGAGCCACGCTGGATTCTTGCTTTGTACAGGAAGATC	5241
Db	275	CCCAAAATTCGGGTTAGCGTGGGAGCCACGCTGGATTCTTGCTTTGTACAGGAAGATC	216
Qy	5242	TACAAGACGACGCAACAGAGTAAAGTGAAGGAAGTTTATTTCAGAAAAATAAGGAGTAT	5301

```
BASE COUNT 143 a 180 c 174 g 188 t 1 others
ORIGIN
Query Match 12.3%; Score 671; DB 14; Length 686;
Best Local Similarity 99.9%; Pred. No. 2e-80;
Matches 671; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4646 AGAACACGAGGCGCCAGCAGATCGGGAAGCTGTTCAGCTTGAGCCAGTCTCTGGACCGGA 4705
DB AGAACACGAGGCGCCAGCAGATCGGGAAGCTGTTCAGCTTGAGCCAGTCTCTGGACCGGA 627
QY 4706 AGTGCAAAAGATGACTCTACCAAGCCTCGAAGAGACCCCATCTCACTCATTTATGGAC 4765
DB AGTGCAAAAGATGACTCTACCAAGCCTCGAAGAGACCCCATCTCACTCATTTATGGAC 567
QY 4766 TCTGGAGAGAGGCGACAAAGAGGACTCTGCCACTCAGCAGCTGAAGTATCAACCA 4825
DB TCTGGAGAGAGGCGACAAAGAGGACTCTGCCACTCAGCAGCTGAAGTATCAACCA 507
QY 4826 GCCCTTGACCTTGAGTCTCGGCTTTGGCTGACCTTCTTTGGGTCTCAGTTTCTTTCTC 4885
DB GCCCTTGACCTTGAGTCTCGGCTTTGGCTGACCTTCTTTGGGTCTCAGTTTCTTTCTC 447
QY 4886 TGCATAA CAAGTTGCCATCTGTTTGCCTTCCAGCACATAAAGTAATGGAACCTTTGATG 4945
DB TGCATAA CAAGTTGCCATCTGTTTGCCTTCCAGCACATAAAGTAATGGAACCTTTGATG 387
QY 4946 CTTTGTCTGGCATTTATGTCTCATGCCAGGATGCCACAGGGGGCCCCAGTCCAGTGG 5005
DB CTTTGTCTGGCATTTATGTCTCATGCCAGGATGCCACAGGGGGCCCCAGTCCAGTGG 327
QY 5006 CTTAACAGCATCTCAGGGAATGCCATCTGGAGCTGGCAAGACCCCTGAGACCTCATAG 5065
DB CTTAACAGCATCTCAGGGAATGCCATCTGGAGCTGGCAAGACCCCTGAGACCTCATAG 267
QY 5066 AGCCTCATCTGTTGGCCACAGCAGCAGCAAGCCTAGAGCCCTCCGGATCCCATCCAGCGCA 5125
DB AGCCTCATCTGTTGGCCACAGCAGCAGCAAGCCTAGAGCCCTCCGGATCCCATCCAGCGCA 207
QY 5126 AAGAGAAATAGAGGAGACATGNAACATTTGCTCTGGCTGTGTACAGGGTGAGCCCA 5185
DB AAGAGAAATAGAGGAGACATGNAACATTTGCTCTGGCTGTGTACAGGGTGAGCCCA 147
QY 5186 AAATTGGGTTTCAAGCTGGAGGAGCCAGCTGGATTCTTGGCTTTGACAGGAAGATCTACA 5245
DB AAATTGGGTTTCAAGCTGGAGGAGCCAGCTGGATTCTTGGCTTTGACAGGAAGATCTACA 87
QY 5246 AAGACCAAGCCACAGAGTAAAGTGAAGGAAGTTTATTACAGAAATAAAGGAGTATACA 5305
DB AAGACCAAGCCACAGAGTAAAGTGAAGGAAGTTTATTACAGAAATAAAGGAGTATACA 27
QY 5306 GCTCTTTTAGAA 5317
DB GCTCTTTTAGAA 15

RESULT 12
BI911134
LOCUS
DEFINITION 603067546F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5216770 5',
mRNA sequence.
ACCESSION BI911134
VERSION BI911134.1 GI:16174696
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 672)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
```

```
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11544 row: k column: 11
High quality sequence stop: 670.
location/Qualifiers
1. 672
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5216770"
/tissue_type="leukocyte"
/lab_host="DH10B"
/clone_lib="NIH_MGC_118"
/note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV
(destroyed); RNA source leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dT primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 027. Note:
this is a NIH_MGC Library."
BASE COUNT 138 a 187 c 189 g 158 t
ORIGIN
Query Match 11.6%; Score 633.6; DB 12; Length 672;
Best Local Similarity 99.1%; Pred. No. 2e-75;
Matches 658; Conservative 0; Mismatches 4; Indels 2; Gaps 2;
QY 3583 GCTTCCCATGTTGCTCAGGCTAAATCTCAAACCTCTCGAAGCTTGGGACTGACGATGACTCTG 3642
DB 1 GCTTCCCATGTTGCTCAGGCTAAATCTCAAACCTCTCGAAGCTTGGGACTGACGATGACTCTG 60
QY 3643 GCTGAGATTGACAGGAAGCTCCCGAGAGTAGTACCGGTGGAACTCTTGGTGGTCT 3702
DB 61 GCTGAGATTGACAGGAAGCTCCCGAGAGTAGTACCGGTGGAACTCTTGGTGGTCT 120
QY 3703 TCTCCT-CCCTCTCAAGGGGACCTGCATACCAAGCCTTTGGGACTGACGATGACTCTG 3761
DB 121 TCTCCTGGCTCTCAAGGGGACCTGCATACCAAGCCTTTGGGACTGACGATGACTCTG 180
QY 3762 GGGCCCCACGGGGCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTTGTACCGAGT 3821
DB 181 GGGCCCCACGGGGCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTTGTACCGAGT 240
QY 3822 TCACCTTCCCTGAGCTGGCTCTACCGCTGGCCCCAACACCGGCTCTGCTTTGTGATGAG 3881
DB 241 TCACCTTCCCTGAGCTGGCTCTACCGCTGGCCCCAACACCGGCTCTGCTTTGTGATGAG 300
QY 3882 AGAAGCGGTGACCGTTGAGATTGAATTTCTGTGTGGGACCCAGTTCTCTGGGTGAGATCAA 3941
DB 301 AGAAGCGGTGACCGTTGAGATTGAATTTCTGTGTGGGACCCAGTTCTCTGGGTGAGATCAA 360
QY 3942 CCCACAGCACAGCTGAGTGTGGCAGGCGCTCTCTCTGGACATCAAGGCTGAGCCCTGGAGC 4001
DB 361 CCCACAGCACAGCTGAGTGTGGCAGGCGCTCTCTCTGGACATCAAGGCTGAGCCCTGGAGC 420
QY 4002 TGTGGAAGCTGACCTCCCTCACTTTGTGGCTCTCCAGGGGGCCCATGTGGACATC 4061
DB 421 TGTGGAAGCTGACCTCCCTCACTTTGTGGCTCTCCAGGGGGCCCATGTGGACATC 480
QY 4062 CTGTTTCCAAATGGCCCACTTTAAAGAGGAGGGGATGCTCTCTGGAGAGCCAGCCAGGCT 4121
DB 481 CTGTTTCCAAATGGCCCACTTTAAAGAGGAGGGGATGCTCTCTGGAGAGCCAGCCAGGCT 539
QY 4122 GGAGCTGCATCAGATAGTTCTGGAAACCCAGCTTCTCCCTTGGAGAGTCTCTCTCTGAA 4181
```

Db	540	GGAGCTGCATCATAGTTCTTGAAACCCAGAGTTCTCCCTTGGGAGTCTCTCTGAA	599
QY	4182	AATGATCCATAATGCCCTCGGTTTCATTCCTCCCTCAGCTCTGTGGTGTGTTTACACCG	4241
Db	600	AATGATCCATAATGCCCTCGGTTTCATTCCTCCCTCAGCTCTGTGGTGTGTTTACACCG	659
QY	4242	CGTC 4245	
Db	660	CGTC 663	
RESULT 13			
LOCUS	BI908442		
DEFINITION	603067392F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5216327 5'	728 bp	linear mRNA EST 16-OCT-2001
ACCESSION	BI908442		mRNA sequence.
VERSION	BI908442.1	GI:16171396	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 728)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgabbs-r@mail.nih.gov		
	Tissue Procurement: Life Technologies, Inc.		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		
	Plate: LLAM11543 row: h column: 24		
	High quality sequence stop: 718.		
FEATURES	Location/Qualifiers		
source	1..728		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5216327"		
	/tissue_type="leukocyte"		
	/lab_host="DH108"		
	/clone_lib="NIH MGC_118"		
	/note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV		
	(destroyed); RNA source leukocytes from anonymous pool of		
	non-activated adult donors. Library is oligo-dT primed		
	and directionally cloned (EcoRV site is destroyed upon		
	cloning). Average insert size 1.7 kb, insert size range		
	1.2-3.3 kb. Library is normalized and enriched for		
	full-length clones and was constructed by C. Gruber		
	(Invitrogen). Research Genetics tracking code 027. Note:		
	this is a NIH MGC Library."		
BASE COUNT	200 a 208 c 200 g 120 t		
ORIGIN			
Query Match	11.6%; Score 633; DB 12; Length 728;		
Best Local Similarity	98.3%; Pred. No. 2.4e-75;		
Matches	713; Conservative 0; Mismatches 5; Indels 7; Gaps 7;		
QY	885	CCATGAATTCGGCGGGTGCACCCAGGCTCAGAGAGAGGGTTTTCAGACAGCTGCC	944
Db	6	CCATGAATTCGGCGGGTGCACCCA-GGCTCAGAGAGAGGGTTTTCAGACAGCTGCC	64
QY	945	TGACACATCTGGACGGCGTGGAGAGAAATCTCTGCTCACTCTCTACCAAGCTCTCC	1004
Db	65	TGACACATCTGGACGGCGTGGAGAGAAATCTCTGCTCACTCTCTACCAAGCTCTCC	124
QY	1005	AAGCTCCCCAGACCATGAGTCTCCAAAGCCAGGAGTCCCAAGCCGCCCATCCACAGC	1064
Db	125	AAGCTCCCCAGACCATGAGTCTCCAAAGCCAGGAGTCCCAAGCCGCCCATCCACAGC	184
QY	1065	AGTGTGGGAGCTGGGATCCCACTCAGCCAGCCTAGCAGCCAGAGAGGAGGC	1124
Db	185	AGTGTGGGAGCTGGGATCCCACTCAGCCAGCCTAGCAGCCAGAGAGGAGGC	244
QY	1125	TCCTGGGACCAATGGCCTCTGGATGAAACGTCAGGAATTTTACTACACAGAAATCAGAG	1184
Db	245	TCCTGGGACCAATGGCCTCTGGATGAAACGTCAGGAATTTTACTACACAGAAATCAGAG	304
QY	1185	AAGAGAGAGAGAGAAATCA-GAGAAAGGAGCCGCCCATGGCAGCGTGGTAGGAAGC	1243
Db	305	AAGAGAGAGAGAGAAATCA-GAGAAAGGAGCCGCCCATGGCAGCGTGGTAGGAAGC	364
QY	1244	CCCCACAGCGCACACAGCCTACAGCCGCCACCCAGCCTGGGAGGCTTCTGTGAGAG	1303
Db	365	CCCCACAGCGCACACAGCCTACAGCCGCCACCCAGCCTGGGAGGCTTCTGTGAGAG	424
QY	1304	AGAGCCTCTGTTCACATGGCCCTGGAAAAATGAGGATTTTAAACCAAAATTCACAGC	1363
Db	425	AGAGCCTCTGTTCACATGGCCCTGGAAAAATGAGGATTTTAAACCAAAATTCACAGC	484
QY	1364	TGCTACTTCTCAAGAGCTCACCCAGAGCCAGATCCCTGCTCAAGAGAGCTGCG	1423
Db	485	TGCTACTTCTCAAGAGATCTCACCCAGAGCCAGATCCCTGCTCAAGAGAGCTGCG	544
QY	1424	CTGA-TTATGTGGA-GGAGAAATCGAGGACATTTAATTCAGATCAGAGCTTATTTGGGCC	1481
Db	545	CTGACTTATGTGGCGGAGAAATCGAGACATTTAATTCAGATCAGAGCTTATTTGG-CC	603
QY	1482	AGGCTTGATACCCAGAACTCTCGCATAGTCATCTCAGGGGGCTGTGGAAATTTGGAA	1541
Db	604	AGGCTTGATACCCAGAACTCTCGCATAGTCATCTCAGGGGGCTGTGGAAATTTGGAA	663
QY	1542	GTCACACTGGCCAGG-CAGGTGAAGAGCTGGGGAGAGCCAGCTGTATGGGACC	1600
Db	664	GTCACACTGGCCAGGCTGAAGAGAGCTGGGGAGAGG-CAGCTGTATGGGACA	722
QY	1601	GGTTC 1605	
Db	723	GGTTC 727	
RESULT 14			
LOCUS	BQ003173		655 bp mRNA linear EST 26-MAR-2002
DEFINITION	UI-H-E11-ayx-n-01-0-UI.s1 NCI_CGAP_E11 Homo sapiens cDNA clone		
ACCESSION	IMAGE:5845056 3', mRNA sequence.		
VERSION	BQ003173		
KEYWORDS	BQ003173.1 GI:19728073		
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 655)		
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),		
JOURNAL	Tumor Gene Index		
COMMENT	Unpublished		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgabbs-r@mail.nih.gov		
	Tissue Procurement: Dr. Jose Mercuende		
	cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa		
	cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa		
	DNA Sequencing by: Dr. M. Bento Soares, University of Iowa		
	Clone Distribution: Clone distribution information can be found		
	through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov		
	The following repetitive elements were found in this cDNA		
	sequence: 32-117, >MER41B#LTR/MER4-group		
	Seq primer: M13 FORWARD		
	POLYA=Yes.		
FEATURES	Location/Qualifiers		


```

source
1. .655
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP E11"
/note="Organ: Left Pelvis; Vector: pT7T3-Pac (Pharmacia)
with a modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI CGAP E11 is a normalized cDNA library containing the
following tissue(s): Chondrosarcoma. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT7T3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
ACACTGGCAC.
TAG_LIB=UI-H-E11
TAG_TISSUE=chondrosarcoma
TAG_SEQ=ACACTGGCAC"
BASE COUNT 143 a 170 c 164 g 178 t
ORIGIN

Query Match 11.5%; Score 628.4; DB 12; Length 655;
Best Local Similarity 99.7%; Pred. No. 1e-74;
Matches 640; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 4676 TGTTAGCTTGAGCCAGTCTGGGACCGGAGTGAAGATGACTCTACCAAGCCCTGA 4735
DB 655 TTTTTCAGCTTGAGCCAGTCTGGGACCGGAGTGAAGATGACTCTACCAAGCCCTGA 596
QY 4736 AGGAGACCCATCTCACTCATTATGGAAGTCTGGGAGAGGCGACAAAAGGACTCC 4795
DB 595 AGGAGACCCATCTCACTCATTATGGAAGTCTGGGAGAGGCGACAAAAGGACTCC 536
QY 4796 TGCCACTCAGCAGCTGAAGTATCAACACAGCCCTTGACCTTGAGTCTGCGCTTGGCT 4855
DB 535 TGCCACTCAGCAGCTGAAGTATCAACACAGCCCTTGACCTTGAGTCTGCGCTTGGCT 476
QY 4856 GACCTCTCTTGGGTCTAGTTTCTTCTCTGCAACAAAGTTCGCATCTGTTTGCCTTC 4915
DB 475 GA-CCTTCTTGGGTCTAGTTTCTTCTCTGCAACAAAGTTCGCATCTGTTTGCCTTC 417
QY 4916 CAGCACTAAAGTAATGGAATTTGATGATGCTTGTGCTGGCATTAATGTCATGCCAG 4975
DB 416 CAGCACTAAAGTAATGGAATTTGATGATGCTTGTGCTGGCATTAATGTCATGCCAG 357
QY 4976 GGATGCCACAGGGGGCCCGAGTCCAGTGGCCCTTAAACAGATCTTCAGGGGAATGTCCATCTG 5035
DB 356 GGATGCCACAGGGGGCCCGAGTCCAGTGGCCCTTAAACAGATCTTCAGGGGAATGTCCATCTG 297
QY 5036 GAGCTGGCAAGACCCCTGAGAGCCTCATAGAGCCTCATCTGTTGGCCACAGAGCCCAAGC 5095
DB 296 GAGCTGGCAAGACCCCTGAGAGCCTCATAGAGCCTCATCTGTTGGCCACAGAGCCCAAGC 237
QY 5096 CTAGAGCCCTCGGATCCCATCCAGCGCAAGAGGAATAGGAGGACATGGAACCATTT 5155
DB 236 CTAGAGCCCTCGGATCCCATCCAGCGCAAGAGGAATAGGAGGACATGGAACCATTT 177
QY 5156 GCCTCTGGGTGTGTACAGGGTGGAGCCCAAAATTTGGGGTTTACGCTGGGAGGCCACCTG 5215
DB 176 GCCTCTGGGTGTGTACAGGGTGGAGCCCAAAATTTGGGGTTTACGCTGGGAGGCCACCTG 117
QY 5216 GATTTCTGGCTTTGTACAGGAAGATCTACAGAGCAAGCCCAACAGAGTAAGTGGAGGA 5275
DB 116 GATTTCTGGCTTTGTACAGGAAGATCTACAGAGCAAGCCCAACAGAGTAAGTGGAGGA 57

BASE COUNT 177 a 180 c 180 g 116 t 2 others
ORIGIN

Query Match 11.5%; Score 623.6; DB 12; Length 655;
Best Local Similarity 98.8%; Pred. No. 4.5e-74;
Matches 648; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

QY 1087 CCACCTCAGCCCGAGCTAGCACCACAGAGAGGAGGCTCTGGGACCCATGGCCTCTG 1146
DB 1 CCACCTCAGCCCGAGCTAGCACCACAGAGAGGAGGCTCTGGGACCCATGGCCTCTG 60
QY 1147 GATGAACGTCAGGAATTTACTACAGAAATCAGAGAAACAGAGAGAGAGAAATCAGAG 1206
DB 61 GATGAACGTCAGGAATTTACTACAGAAATCAGAGAAACAGAGAGAGAGAAATCAGAG 120

1. .655
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="TCBAP9080"
/tissue_type="leukemia"
/lab_host="DH10B"
/note="vector: lambda pSB; Site 1: BamHI; Site 2: EcoRI;
First strand cDNA was primed with an anchored
XhoI-oligo(dT) primer [5'-GGAGGACTCGAGCGCCGAGGAGGAG(T)VN
3'; V=A,C,G; N=A,C,G,T] and then dg tailed. Second strand
was primed with a BamHI-dC primer
[5'-AGAGAGCTCGGATCGCGCGCGCAATAATAATAAT(C) 3'].
Double-stranded cDNA was then digested with BamHI and XhoI
and directionally cloned into the BamHI and SalI sites of
lambda pSB vector. Library was constructed by Wei Yu at RIKEN
normalization. Library was constructed by Wei Yu at RIKEN
of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,
Itoh M, Nagaoka S, Sasaki N, Okazaki Y, Muramatsu M,
Schneider C, Hayashizaki Y, High efficiency selection of
full-length cDNA by improved biotinylated cap trapper.,
DNA Res 4: 1, 61-6, Feb 28, 1997)"
BASE COUNT 177 a 180 c 180 g 116 t 2 others
ORIGIN

Query Match 11.5%; Score 623.6; DB 12; Length 655;
Best Local Similarity 98.8%; Pred. No. 4.5e-74;
Matches 648; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

QY 1087 CCACCTCAGCCCGAGCTAGCACCACAGAGAGGAGGCTCTGGGACCCATGGCCTCTG 1146
DB 1 CCACCTCAGCCCGAGCTAGCACCACAGAGAGGAGGCTCTGGGACCCATGGCCTCTG 60
QY 1147 GATGAACGTCAGGAATTTACTACAGAAATCAGAGAAACAGAGAGAGAGAAATCAGAG 1206
DB 61 GATGAACGTCAGGAATTTACTACAGAAATCAGAGAAACAGAGAGAGAGAAATCAGAG 120

```

QY	1207	AAAGGCAGGCCCCCATGGCAGCGGTGGTAGGAACGCCCCACAGGCGCACACCGCCTA	1266
Db	121	ATTGGCAGGCCCCCATGGGCAGCGGTGGTAGGAACGCCCCACAGGCGCACACCGCCTA	180
QY	1267	CAGCCCCACACACCCATGGGAGCCTTCTGTGAGAGAGGCTCTGTTCCACATGGCCC	1326
Db	181	CAGCCCCACACACCCATGGGAGCCTTCTGTGAGAGAGGCTCTGTTCCACATGGCCC	240
QY	1327	TGGAAAAATGAGGATTTTAACCAAAAATTCACACAGCTGCTACTTCTACAAAGACCTCAC	1386
Db	241	TGGAAAAATGAGGATTTTAACCAAAAATTCACACAGCTGCTACTTCTACAAAGACCTCAC	300
QY	1387	CCCAGAAGCCAAGATCCCTGGTCAAGAGAAGCTGGCCTGATTATGTGGAGGAGATCGA	1446
Db	301	CCCAGAAGCCAAGATCCCTGGTCAAGAGAAGCTGGCCTGATTATGTGGAGGAGATCGA	360
QY	1447	GGACATTTTAATTGAGATCAGAGACTTATTGGCCAGGCTGGATACCCCAAGACCTCGC	1506
Db	361	GGACATTTTAATTGAGATCAGAGACTTATTGGCCAGGCTGGATACCCCAAGACCTCGC	420
QY	1507	ATAGTCATACTGCAGGGGCTGCTGGAATTGGGAAGTCAACACTGGCCAGGCGAGGTGAAG	1566
Db	421	ATAGTCATACTGCAGGGGCTGCTGGAATTGGGAAGTCAACACTGGCCAGGCGAGGTGAAG	480
QY	1567	GAAGCCTGGGGAGAGGCCAGCTGTATGGGGACCGCTTCCAGCATGCTTCTACTTTCAGC	1626
Db	481	NAAGCCTGGGGAGAGGCCAGCTGTATGGGGACCGCTTCCAGCATGCTTCTACTTTCAGC	540
QY	1627	TGCAGAGAGTGGCCCGAGTCCAGGTGGTGAAGTCTCGCTGAGCTATCGGAAAAGATGGG	1686
Db	541	TGCAGAGAGTGGCCCGAGTCCAGGTGGTGAAGTCTCGCTGAGCTATCGG-AAAAGATGGG	599
QY	1687	ACAGCCACTCCGGCTCCCATTAGACAGATCCTGTCTAGGCCAGGCGGCTGCTTT	1742
Db	600	ACAGCCACTCCGGCT-CCATTAGACAGATCCTGTCTAGGCCAGGCGGCTGCTTT	654

Search completed: January 29, 2004, 13:05:49
 Job time : 10855.3 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic. - nucleic search, using sw model

Run on: January 29, 2004, 03:02:31 ; Search time 291.391 Seconds
(without alignments)
8246.264 Million cell updates/sec

Title: US-09-996-617-1
Perfect score: 5444
Sequence: 1 gccccaggctggagaggt.....aggataagaagttacctac 5444

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents,NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	165.8	3.0	1371	2	US-08-910-731-1
2	165.8	3.0	1371	2	US-08-795-395-1
3	163.6	3.0	1371	2	US-08-910-731-7
4	163.6	3.0	1386	2	US-08-910-731-5
5	156.2	2.9	1374	2	US-08-910-731-3
6	156.2	2.9	1374	2	US-08-795-395-3
7	115	2.1	585	4	US-09-340-620A-50
8	115	2.1	740	4	US-09-340-620A-48
9	106.2	2.0	579	4	US-09-340-620A-62
10	106.2	2.0	777	4	US-09-340-620A-60
11	62	1.1	7218	1	US-08-232-463-14
12	59.4	1.1	257	4	US-09-016-434-208
13	57.2	1.1	3116	4	US-09-904-615-43
14	49	0.9	337	1	US-08-594-031-79
15	48.6	0.9	1878	4	US-09-465-558-39
16	48	0.9	1582	3	US-08-545-196B-10
17	48	0.9	1582	3	US-08-545-196B-12
18	47.2	0.9	17949	3	US-09-087-465-3
19	46.2	0.8	599	3	US-09-328-111-147
20	46.2	0.8	1441	3	US-08-821-994-63
21	46.2	0.8	4141	4	US-09-245-281-42
22	46.2	0.8	4141	4	US-09-207-359B-42
23	46.2	0.8	4141	4	US-09-340-620A-42
24	45.8	0.8	924	1	US-08-468-709B-1
25	45.8	0.8	924	1	US-08-241-664B-1
26	45.8	0.8	924	5	PCT-US93-03936-1
27	45.8	0.8	1717	1	US-08-468-709B-6

Sequence 6, Appli
Sequence 6, Appli
Sequence 14, Appli
Sequence 17, Appli
Sequence 17, Appli
Sequence 1, Appli
Sequence 3, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 18, Appli
Sequence 1, Appli
Sequence 26, Appli
Sequence 17, Appli
Sequence 89, Appli
Sequence 6, Appli
Sequence 166, Appli
Sequence 4, Appli

US-08-241-664B-6
PCT-US93-03936-6
US-08-232-463-14
US-09-007-005-17
US-09-244-796-17
US-09-039-046-1
US-09-734-250-3
US-09-390-207-1
US-08-234-245-1
US-08-841-349-18
US-08-658-883B-1
US-09-676-610B-26
US-08-860-339-17
US-09-489-847-89
US-08-300-903A-6
US-08-988-197-6
US-09-220-132-166
US-08-897-340-4

ALIGNMENTS

RESULT 1
US-08-910-731-1
; Sequence 1, Application US/08910731
; Patent No. 5932440
; GENERAL INFORMATION:
; APPLICANT: CHATTERJEE, DEB K.
; APPLICANT: SHANDILYA, HARINI
; TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,731
; FILING DATE: (Herewith)
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/795,395
; FILING DATE: 04-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/794,546
; FILING DATE: 03-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,057
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.3440003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1371 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS

i	LOCATION:	1...1368	
US-08-910-731-1			
Query Match	3.0%;	Score 165.8; DB 2;	Length 1371;
Best Local Similarity	56.0%;	Pred. No. 8,7e-34;	
Matches 314; Conservative	0;	Mismatches 247;	Indels 0; Gaps 0;
QY	2920	CTCTTTCCCTCCTCAAGGTCACAGAAACTGAAGGAGCTGGACCTTAATGTGAAACTCG	2979
Db	811	CTCTGCCGTCTCCTCCAGGCACAAGGAGACCCTGAAGGAGCTCAGTCTGGCGGGCAACAAG	870
QY	2980	CTGAGCCACTCTGCAGTGAAGAGTCTTTGTAAAGACCCTGAGAGCCCTCCCTGCCTCCCTG	3039
Db	871	CTGGGGCACAGGGCGCCCGGCTGTGTGCGAGAGCCTGTGTCAGCCCGGCTGCCAGCTG	930
QY	3040	GAGACCCCTGGGTTGGCTGTGGCGCTCACAGCTGAGGAGCTGCAAGGACCTTGCTGCTTT	3099
Db	931	GAGTCCCTGTGGGTGAAGTCTTCAGCGCTCAGCGCGGCTGCTGCCAGCAGCTCAGCTTG	990
QY	3100	GGGCTGAGAGCCAACAGACCTTGACCGAGCTGGACCTGAGTCTTAATGTCTACGGAT	3159
Db	991	ATGCTTACCAGAACAGCATCTCTTGGAACTTCAGTTGAGCAGCAACAAGCTGGGTGAC	1050
QY	3160	GCTGGAGCCAAAACACTTTGGCCAGAGACTCAGACAGCCGAGCTGCAAGCTACAGCAGCTG	3219
Db	1051	TCTGGCATCCAGAGCTGTGCCAGGCCCTGAGCGACCGGGACCAACTGCGGGTGCTC	1110
QY	3220	CAGCTGTCAGCTGTGGCCTCAGCTGTGACTGTGTCGAGGACCTGGCCTCTGTGCTTAGT	3279
Db	1111	TGTCCTTTGGGACTGTGAGGTGACCAAAGCGGCTGCAGCAGCCTCGCTCGCTCTGCTG	1170
QY	3280	GCCAGCCCCAGCTGAAGGAGCTTAGACCTGCAGAGAACAACTGAGATGAGCTTGGCGTG	3339
Db	1171	GCCAAACCGAGCCTGGCAGAGCTGGACCTGAGCAAACAATGTGTGGCGCACCCGGGGCTC	1230
QY	3340	CGACTGCTGTGAGGGGCTCAGGCATCTGCTGCTGCAAACTCATAGCCTGGGGCTGGAC	3399
Db	1231	CTCAGCTGTGGGAGCCTTGGAGCAGCCGGGCTGGCCCTTGAGCAGCTGTGTCTGTATC	1290
QY	3400	CAGACAACTCTGAGTGATGAGATGAGCAGGAACCTGAGGGCCCTGGAGCAGGAGAAACCT	3459
Db	1291	GACACCTACTGGACGGAGAGGTGGAGGACCGGCTGCAGGCCCTTGAGGGGAGCAAGCCC	1350
QY	3460	CAGCTGCTATCTTCAGCAGA	3480
Db	1351	GGCCTGAGGCTATCTCCTGA	1371

825	TCTGTGCGCTGTCTCTCAGGCGCAAGGAGACCTGAAAGGAGCTCAGCTGCGCCGGCAACGA	884
2979	GCTGAGCCACTCTGCAGTGAAGAGTCTTTTGAAGACCCTGAGACGCCCTCGCTGCCTCCT	3038
885	GCTGGGGATGAGGGTGCSCCGACTGTGTGTGTGAGACCTCTGTGAACTCTGGCTGCCAGCT	944
3039	GGAGACCTCTGGGTTGGCTGGCTGTGGCTCTCAGCTGAGGACTGTGCAAGGACCTTGGCTTT	3098
945	GGAGTGCCTGTGGGTGAAGTCTCTGCAGCTTTCAGACGCCCTGTCTGCCCCCACTTCAGCTC	1004
3099	TGGGCTTGAGAGCCAAACAGACCCCTGACCGAGCTGGACTCAGCTTCAATGTGTCTCAGGA	3158
1005	AGTGTCTGGCCAGAAACAGTTTCTCTGGAGCTACAGATTAAGCAAACAACAGGCTGGAGGA	1064
3159	TGCTGGAGCCAAACACCTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCGACT	3218
1065	TGCGGGCTGCGGAGCTGTGCCAGGCCCTGGCCAGCCTGGCTCTGTGCTGGGGTGCT	1124
3219	GCAGCTGGTCAGCTGTGGCTCTCAGTCTGACTGCTGCAGGAACCTGGCTCTCTGTGCTTAG	3278
1125	CTGGTTGGCCGACTGCGATGTGATGACAGCAGCTGCAGCAGCTGCGCCAAACCTGTT	1184
3279	TGCCAGCCCGAGCCTGAAGAGACTAGACTTGCAGCAGAAACAACCTGGATGACGTTGGCGT	3338
1185	GGGCAACCAAGGCTTGGTGTGGAGCTTCACTCAGCAAACAATGCTGGGGAGCCCGGCAT	1244
3339	GGGACTGCTCTGTGAGGGGCTCAGGCATCTGCTGCAAACTCATACGCTGGGGCTGGA	3398
1245	CCTGCACTGGTGGAGAGCGTTCGGCAGCCGGGCTGCCTCCTGGAGCAGCTGGTCTGTGA	1304
3399	CCAGAACACTCTGAGTGAATGAGATGAGGCAAGGAACTGAGGGCCCTGGAGCAGGAGAAACC	3458
1305	CGACATTTACTGGTCTGAGGAGATGGAGGACCGGCTGCAGGCCCTGGAGAGGACAGACC	1364
3459	TCAGCTGCTCATCTTCAGCAGA	3480
1365	ATCCCTGAGGCTCATCTCTGA	1386

RESULT 5

```

US-08-910-731-3
; Sequence 3, Application US/08910731
; Patent No. 5932440
; GENERAL INFORMATION:
; APPLICANT: CHATTERJEE, DEB K.
; APPLICANT: SHANDILYA, HARINI
; TITLE OF INVENTION: Mamalian Ribonuclease Inhibitors and Use Thereof
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910.731
; FILING DATE: (Herewith)
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/795,395
; FILING DATE: 04-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/794,546
; FILING DATE: 03-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,057
; FILING DATE: 16-AUG-1996

```

ATTORNEY/AGENT INFORMATION:
NAME: ESMOND, ROBERT W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.3440003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1374 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1368
US-08-910-731-3

Query Match 2.9%; Score 156.2; DB 2; Length 1374;
Best Local Similarity 54.2%; Pred.No.3.1e-31;
Matches 317; Conservative 0; Mismatches 268; Indels 0; Gaps 0;

QY	2896	GTCA	CAGATGCCTATTGGCAGATTCTTCTCGTCTCAAGGTACACGAAACCTGGAAG	2955
DB	787	GTCA	CTGCAGAAAGGCTGCAAGGACCTGTGCGGTCTCTCAGAGCCAAAGCAGACCTGGAAG	846
QY	2956	GAGT	CGACTAAGTGGAAATCGCTGAGCCACTCTGCAGTGAAGAGTCTTTGTAAAGACC	3015
DB	847	GAACT	CAGCTAGCTGGCAATGAGCTGAAGATGAGGGTGCCTCACTGCTGTGTGAGAGC	906
QY	3016	CTGA	CAGCGCCCTCGCTGCCCTCTCTGGAGACCCCTGCGGTGTGGCTGTGGCTTCACAGCT	3075
DB	907	CTGT	TAGAGCTGGCTGTGAGTGGAGTCACTGTGGGTAAAGACCTGTAGCCTTCACAGCT	966
QY	3076	GAGG	ACTGCAAGCACTTGCCTTTGGGCTGAGAGCCAAACAGACCTTCAGCGAGCTGGAC	3135
DB	967	GCTC	TTGTCCTCCTCTGCTCGGTGTTGACCAAAAACAGTTCTCTGTTGAGTTGCAA	1026
QY	3136	CTGA	GCTCAATGTCTACGGATGCTGGAGCCAAACACCTTTGCCAGAGACTTGAGACAG	3195
DB	1027	ATGA	CAGAGCAAAACCGCTGGGAGACTCGGAGTCTGGGAGCTTTGCAAGCCCTGGGCTAT	1086
QY	3196	CCGAG	CTGCAAGCTACAGCGACTGCGAGCTGGTCTGAGCTGGGCTCACTGCTGAGCTGCTGC	3255
DB	1087	CCGA	CACAGTGTGCTGTGCTTTGGCTGGGAGACTGTGATGTGACAGACAGTGGCTGC	1146
QY	3256	CAGG	ACTGCTGTGTTAGTGCACGCCACGCTGAAAGAGCTAGACCTGACAGCAG	3315
DB	1147	AGCA	GCTTGGCACTGTCTCTGCTGGCCAAACCGAGCTTGAGGGAACTGAGACCTCAGTAAC	1206
QY	3316	ACA	ACTGATGACGTTGGCTGGAGTCTGCTGTGAGGGGCTCAGGATCTCGCTGCTGC	3375
DB	1207	AACT	GATGGGGACAAACGGTGTCTTACAACTGCTGGAGAGCCTCAAAACAGCCAGCTGC	1266
QY	3376	AAAC	TATACGCTGGGCTGGACAGACAACTCTGAGTGTGAGATGAGGAGGAACTG	3435
DB	1267	ATCC	TTACAGAGTTGCTGTATGACATTTACTGACGATGAGTGGAAACACAGCTT	1326
QY	3436	AGG	GGCCTGGAGCAGGAAACCTCAGCTGCTCATCTTCAGCAGA	3480
DB	1327	CGGG	CCCTGGAGGAAAGGCCATCCCTGAGGATCATTTCTCTGA	1371

RESULT 6
US-08-795-395-3
Sequence 3, Application US/08795395
Patent No. 5965399
GENERAL INFORMATION:
APPLICANT: CHATTERJEE, DEB K.
TITLE OF INVENTION: Cloning and Expression of Rat Liver and Porcine Liver Ribonuclease Inhibitor

NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
 STREET: 1100 NEW YORK AVE., N.W., SUITE 600
 CITY: WASHINGTON
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20005-3934
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/795.395
 FILING DATE: 04-FEB-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/024,057
 FILING DATE: 16-AUG-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: ESMOND, ROBERT W.
 REGISTRATION NUMBER: 32,893
 REFERENCE/DOCKET NUMBER: 0942.3440002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-371-2600
 TELEFAX: 202-371-2540
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1374 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: both
 TOPOLOGY: both
 MOLSCULE TYPE: CDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1368
 US-08-795-395-3

Query Match	2.9%	Score 156.2;	DB 2;	Length 1374;
Best Local Similarity	54.2%;	Pred. No. 3.1e-31;		
Matches 317;	Conservative	0;	Mismatches 268;	Indels 0; Gaps 0;
2896	GTCAAGATGCTATTGGCAGATTCTTCTCCGTCTCTCAAGTGCACAGAAACCTTGAAG	2955		
Db				
787	GTCACTTGCAGAAAGCTGCAGAGACCTGTGCCGTGTCTTCAGAGCAAGCAGACCTTGAAG	846		
QY				
2956	GAGCTGGACCTAAAGTGGAAATCTCGCTGAGCCACTCTGCAGTGAAGAGTCTTTTGAAGACC	3015		
Db				
847	GAACTCAGCCTAGCTGGCAATGAGCTGAAGGATGAGGGTGCCCAACTGCTGTGTGAGAGC	906		
QY				
3016	CTGAGAGCCCTCGCTGCTCTCTTGGAGACCCCTGGGTTGGCTGGCTGTGGCTTCCACAGCT	3075		
Db				
907	CTGTTAAGCCTTGGCTGTCAAGTGGAGTCACTGTGGGTAAAGACCTGTGAGCTTCAAGCT	966		
QY				
3076	GAGGACTGCAAGACCTTGGCTTTGGGCTGAGAGCCAAACAGACCCCTGCAGCGAGCTGGAC	3135		
Db				
967	GCCTCTGTCCCCACTTCTGCTCGGTTGTGACAAACAGATTCTCTGTTTGAGTTGCAA	1026		
QY				
3136	CTGAGCTTCAATGTGCTACCGGATGTGTGGAGCCAAACACCTTTGCCAGAGACTTGAGACAG	3195		
Db				
1027	ATGAGCAGCAACCCGCTGGGAGACTCGGAGTCTGTGGAGCTTTTGCAAGGCCCTGGGCTAT	1086		
QY				
3196	CCGAGCTGAAGCTACAGCAGACTGCAGCTGGTCAGCTGTGGCCCTCAGCTCTGACTGCTGC	3255		
Db				
1087	CCGGACACAGTGTGCGTGTGCTTTGGCTGGGAGACTGTGATGTGACAGACAGTGGCGTCG	1146		
QY				
3256	CAGGACCTGGCCTCTGTGCTTAGTGGCAGCCCGCAGCCTGAAGGAGCTAGACCTTGCAGCAG	3315		
Db				
1147	AGCAGCCTTGCCACTGTCTTGTCTGGCCCAACCGCAGCTTTGAGGNACTGGACCTTCACTAAC	1206		
QY				
3316	AACAACCTGGATGACGTTGGCGTGCAGTCTCTGTGAGGGGGCTCAGGCCATCCTGCTGCTC	3375		
Db				

Db	1207	A	CTGCATGGGGACAACGCTGTCTTACAAC	TCTGGAGAGCCTCAACAGGCCAGCTGC	1266
Qy	3376	A	ACTATACGCTTGGGCTGGACACGACAACTCT	GATGATGAGATGAGGCAGGAAC	3435
Db	1267	A	TCTTCAGCAGCTGTCTCTGTATGATTTACT	TGGACCGATGAGGTGGAAGACCACTT	1326
Qy	3436	A	GGCCCTTGGAGCAGGAGAACTCAGTCTGCTCAT	CTTTCAGCAGA	3480
Db	1327	C	GGGCTTGGAGAGAAAGGCCATCCTT	CAGGATCATTTCTTGA	1371

RESULT 7

```

US-09-340-620A-50
; Sequence 50, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340, 620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 585
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-340-620A-50

```

Query Match	2.1%;	Score 115;	DB 4;	Length 585;
Best Local Similarity	61.5%;	Pred. No. 1.5e-20;		
Matches 184;	Conservative 0;	Mismatches 115;	Indels 0;	Gaps 0;
QY	4476	CCCTCCAGCCCGATAGCCGTACCTTCACCTCTGGATGCCCGCAGCTTGCTGCACATTGT	4535	
Db	285	CGGCCAGCTGGGATCCAGGCCCTCTCCTAGTCGGCAGCCAAAGCCAGGCTGTCACATTAT	344	
QY	4536	GGACCAGTATCGAGAGCAGCTGATAGCCCGCAGGTGACATCGTGTGGAGTTCTCTTGGACAA	4595	
Db	345	AGACCAACACCGGGCTCGCTTTATCGCGAGGGGTACAAAGCTTGAGTGGGTGCTGTGATGC	404	
QY	4596	ACTGCATGGACAGGTGCTGAGCCAGGAGCAGTAGTACGAGAGGGTCTGGCTCAGAAACACGAG	4655	
Db	405	TCTGTACGGGAAGGTCTTGCAGATGAGCAGTACCAGGCAGTGGCGCCGAGCCCA	464	
QY	4656	GCCCAGCCAGATCGGAAGCTGTTACGCTTTGAGCCAGTCTCTGGGACCGGAAGTGCAAGA	4715	
Db	465	CCCAAGCAAGATCGGAAGCTCTTTCAGTTTTCACACACAGCCTGGAACTGGACCTGCAAGGA	524	
QY	4716	TGGACTTCAACAAGCCCTGAAGGAGACCCATTCCTCACTCATTTATGGAACCTCTGGGAGA	4774	
Db	525	CTTGCTCTCTCAGGCCCTAAGGAGTCCCAAGTCTCTACTGTTGGAGACCTGGAGCGGA	583	

RESULT 8

RESUB. 8
 US-09-340-620A-48
 ; Sequence 48, Application US/09340620A
 ; Patent No. 6482933
 ; GENERAL INFORMATION:
 ; APPLICANT: Bertin, John
 ; TITLE OF INVENTION: NOVEL MOLECULES OF THE
 ; FILE REFERENCE: 07334-124001
 ; CURRENT APPLICATION NUMBER: US/09/340,620A
 ; CURRENT FILING DATE: 1999-06-28
 ; PRIOR APPLICATION NUMBER: US 09/245,281

WARD-RELATED PROTEIN FAMILY AND USES THERE

```

; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (34)...(638)
; US-09-340-620A-48

Query Match      2.1%; Score 115; DB 4; Length 740;
Best Local Similarity 60.9%; Pred. No. 1.8e-20;
Matches 187; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 4476 CCTCCAGCCCGCATACCGTACCTTACCTCTGGATGCCCGCAGTTCCTGCACTTTGT 4535
DB 338 CGGCCAGCTGGGATCCAGSCCCTCTCAGTCGGCAGCAAGCCAGGCGCTGCACTTTAT 397
QY 4536 GGACCAAGTATCAGACGACGTATAGCCCGAGTGACATCGGTGGAGTTGTCTTGGACAA 4595
DB 398 AGACCAAGCACCAGGCTGCGCTTATCCGAGGGGTACAAAAGTGTGAGTGGCTCTGGATGC 457
QY 4596 ACTGCATGTGACAGGTCTGAGCCAGGACGAGTACGAGAGGGTGTGGCTGAGAACACCGAG 4655
DB 458 TCTGTACGGGAAGTCTCTACGGATGAGCAGTACGAGGAGTGGCGGCCGAGCCACCA 517
QY 4656 GCCAGCCAGATGCCGAAGCTGTTCAGCTTGAGCCAGTCTTGGACCGGAAGTGCAGAA 4715
DB 518 CCCAAGCAAGATGCCGAAGCTTTCAGTTTACACCCAGGCTGGAACCTGGACCTGCAAGGA 577
QY 4716 TGGACTCTACCAAGCCCTGAAGAGAGCCATCTCCTACCTTATGGAACCTGGAGAA 4775
DB 578 CTGTCTCTCCAGGCCCTTAAAGGAGTCCAGCTCTACCTGCTGGTGGAGACCTGGAGCGGAG 637
QY 4776 GGGCAGC 4782
DB 638 CTGAGGC 644

RESULT 9
US-09-340-620A-62
; Sequence 62, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 62
; LENGTH: 579
; TYPE: DNA
; ORGANISM: Mus musculus
; US-09-340-620A-62

```

```

Query Match      2.0%; Score 106.2; DB 4; Length 579;
Best Local Similarity 64.4%; Pred. No. 3.3e-18;
Matches 159; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 4528 CACTTTGTGGACCAAGTATCGAGAGCAGCTGATAGCCCGAGTGACATCGGTGGAGTTGTC 4587
DB 331 CACTTTGTGGACCAAGCAGCAAGCAAGCACTCATTTCCAGGGTTCACAGAAAGTGGACGAGTG 390
QY 4588 TTGGACAAACTGCATGGACAGGTGCTGAGCCAGGACAGTACGAGAGGGTGTGGCTGAG 4647
DB 391 CTGGATGCTTTTGCATGGCAGTGTCTGACTGAGGACAGTACCGAGGAGTTCGTGCAGAG 450
QY 4648 AACACGAGGCCCGACGACATCGGGAAGCTGTTCAGCTTTGAGCCAGTCTCTGGACCGGAAG 4707
DB 451 ACCACCAAGCAAGATGAGGAGCTCTTCAGCTTTGTTCCATCTCTGGAACCTGACC 510
QY 4708 TGCAAAGATGGAAGTCTTACCAAGCCCTGAAAGAGAGCCCATCTCTACCTCATTTATGGAATC 4767
DB 511 TGCAGGAGTCCCTCTCCAGGCTTGAAGGAAATACATCCCTACTTTGGTGTATGACCTG 570
QY 4768 TGGGAGA 4774
DB 571 GAGCAGA 577

RESULT 10
US-09-340-620A-60
; Sequence 60, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (89)...(667)
; US-09-340-620A-60

Query Match      2.0%; Score 106.2; DB 4; Length 777;
Best Local Similarity 64.4%; Pred. No. 4e-18;
Matches 159; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 4528 CACTTTGTGGACCAAGTATCGAGAGCAGCTGATAGCCCGAGTGACATCGGTGGAGTTGTC 4587
DB 419 CACTTTGTGGACCAAGCAGCAAGCAAGCACTCATTTCCAGGGTTCACAGAAAGTGGACGAGTG 478
QY 4588 TTGGACAAACTGCATGGACAGGTGCTGAGCCAGGACAGTACGAGAGGGTGTGGCTGAG 4647
DB 479 CTGGATGCTTTTGCATGGCAGTGTCTGACTGAAAGACAGTACCGAGGAGTTCGTGCAGAG 538
QY 4648 AACACGAGGCCCGACGACATCGGGAAGCTGTTCAGCTTTGAGCCAGTCTCTGGACCGGAAG 4707
DB 539 ACCACCAAGCAAGATGAGGAGCTCTTCAGCTTTGTTCCATCTCTGGAACCTGACC 598
QY 4708 TGCAAAGATGGAAGTCTTACCAAGCCCTGAAAGAGAGCCCATCTCTACCTCATTTATGGAATC 4767
DB 599 TGCAGGAGTCCCTCTCCAGGCTTGAAGGAAATACATCCCTACTTTGGTGTATGACCTG 658

```

QY	4768	TGGGAGA	4774
Db	659	GAGCAGA	665
RESULT 11			
QY	US-08-232-463-14	Application US/08232463	
Db	Sequence 14	Patent No. 5670367	
	GENERAL INFORMATION:		
	APPLICANT: DORNER, F.		
	APPLICANT: SCHEIFLINGER, F.		
	APPLICANT: FALKNER, F. G.		
	TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS		
	NUMBER OF SEQUENCES: 52		
	CORRESPONDENCE ADDRESS:		
	ADDRESSEE: Foley & Lardner		
	STREET: 1800 Diagonal Road, Suite 500		
	CITY: Alexandria		
	STATE: VA		
	COUNTRY: USA		
	ZIP: 22313-0299		
	COMPUTER READABLE FORM:		
	MEDIUM TYPE: Floppy disk		
	COMPUTER: IBM PC compatible		
	OPERATING SYSTEM: PC-DOS/MS-DOS		
	SOFTWARE: Patent in Release #1.0, Version #1.25		
	CURRENT APPLICATION DATA:		
	APPLICATION NUMBER: US/08/232.463		
	FILING DATE:		
	CLASSIFICATION: 435		
	PRIOR APPLICATION DATA:		
	APPLICATION NUMBER: US/07/935.313		
	FILING DATE:		
	APPLICATION NUMBER: EP 91 114 300.6		
	FILING DATE: 26-AUG-1991		
	ATTORNEY/AGENT INFORMATION:		
	NAME: BENT, Stephen A.		
	REGISTRATION NUMBER: 29,768		
	REFERENCE/DOCKET NUMBER: 30472/114 IMMU		
	TELECOMMUNICATION INFORMATION:		
	TELEPHONE: (703)836-9300		
	TELEFAX: (703)683-4109		
	TELEX: 899149		
	INFORMATION FOR SEQ ID NO: 14:		
	SEQUENCE CHARACTERISTICS:		
	LENGTH: 7218 base pairs		
	TYPE: nucleic acid		
	STRANDEDNESS: single		
	TOPOLOGY: linear		
	IMMEDIATE SOURCE:		
	CLONE: pTZapt-F1s		
	US-08-232-463-14		
	Query Match	1.1%;	Score 62; DB 1; Length 7218;
	Best Local Similarity	2.8%;	Pred. No. 8.5e-06;
	Matches	11; Conservative	230; Mismatches 145; Indels 0; Gaps 0;
QY	23	GAAGAAACCTGGAGCCAGCAGCCCGGGCTCCACTCTGGGTTCTGAAAGCCCATTCCT	82
Db	1052	GAGGAGCTTGCATYY	1111
QY	83	GCTCTGGCGCTCTCCACCCACCTCTTCTCAGCTTGCAGCTCAGGTTGATCTCAG	142
Db	1112	YYY	1171
QY	143	GAGTCCAGGACCCAGGAGGAGGAAGATCTGAGGAACACAGACAGTGAGGTTGCCAC	202
Db	1172	YYY	1231
QY	203	ACCCCATCTCCCGTACACCATCTCCCTCCCTCACCTCCCTCCCTCCCTCCCTCCCT	262

QY	4768	TGGGAGA 4774
Db	659	GAGCAGA 665
RESULT 11		
QY	US-08-232-463-14	Application US/08232463
Db	Patent No. 5670367	GENERAL INFORMATION:
	APPLICANT: DORNER, F.	APPLICANT: SCHEIFLINGER, F.
	APPLICANT: FALKNER, F. G.	APPLICANT: FALKNER, F. G.
	TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS	
	NUMBER OF SEQUENCES: 52	
	CORRESPONDENCE ADDRESS:	
	ADDRESSEE: Foley & Lardner	
	STREET: 1800 Diagonal Road, Suite 500	
	CITY: Alexandria	
	STATE: VA	
	COUNTRY: USA	
	ZIP: 22313-0299	
	COMPUTER READABLE FORM:	
	MEDIUM TYPE: Floppy disk	
	COMPUTER: IBM PC compatible	
	OPERATING SYSTEM: PC-DOS/MS-DOS	
	SOFTWARE: Patent in Release #1.0, Version #1.25	
	CURRENT APPLICATION DATA:	
	APPLICATION NUMBER: US/08/232.463	
	FILING DATE:	
	CLASSIFICATION: 435	
	PRIOR APPLICATION DATA:	
	APPLICATION NUMBER: US/07/935.313	
	FILING DATE:	
	APPLICATION NUMBER: EP 91 114 300.6	
	FILING DATE: 26-AUG-1991	
	ATTORNEY/AGENT INFORMATION:	
	NAME: BENT, Stephen A.	
	REGISTRATION NUMBER: 29,768	
	REFERENCE/DOCKET NUMBER: 30472/114 IMMU	
	TELECOMMUNICATION INFORMATION:	
	TELEPHONE: (703)836-9300	
	TELEFAX: (703)683-4109	
	TELEX: 899149	
	INFORMATION FOR SEQ ID NO: 14:	
	SEQUENCE CHARACTERISTICS:	
	LENGTH: 7218 base pairs	
	TYPE: nucleic acid	
	STRANDEDNESS: single	
	TOPOLOGY: linear	
	IMMEDIATE SOURCE:	
	CLONE: PT2gpt-F1s	
	US-08-232-463-14	
Query Match 1.1%; Score 62; DB 1; Length 7218;		
Best Local Similarity 2.8%; Pred. No. 8.5e-06;		
Matches 11; Conservative 230; Mismatches 145; Indels 0; Gaps 0;		
QY	23	GAAGAAACCTGGAGCCAGCAGCCCGGGCTCCACTCTGGGTTCTGAAAGCCCATTCCT 82
Db	1052	GAGGAGCTTGGATYY 1111
QY	83	GCTCTGGCGCTCTCCACCCACCTCTTCTCAGCTTGCAGCTCAGGTTGATCTCAG 142
Db	1112	YYY 1171
QY	143	GAGTCCAGGACCCAGGAGGGAAGATCTGAGGAACACAGACAGTGAGGTTGCCAC 202
Db	1172	YYY 1231
QY	203	ACCCCATCTCCCGTACACCATCTCCCTCCCTCACCTCCCTCCCTCCCTCCCTG 262


```

; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-465-558-39

Query Match      0.9%; Score 48.6; DB 4; Length 1878;
Best Local Similarity 52.8%; Pred. No. 0.013;
Matches 105; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 5237 AGATCTACAAGAGCAAGCCACAGAGTAAAGTGGAGGAAGTTTATTTCAGAAAAATAAGG 5296
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1672 AGATCCATGATGNAGCGCCACCATATGTAAGAGGATCCAGCATTTTACAGATATGGG 1731

QY 5297 AGTATCACAGCTCTTTTAGAATTTGTCTAGCAGGCTTCCAGTTTTCACGAAACCCC 5356
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1732 AACTTTGTCGATAGTTTCTTATTTCAGGCGACATACCTGTAAGATGCTTCGCTGATATGCTA 1791

QY 5357 TATAAATTAAAAATTTTTTACTTAAATTTAAGAAATTAAGAAATCAAAAAAGAAAAAT 5416
    || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1792 TAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1851

QY 5417 GAAATATAAGGAATAAGAA 5435
    || | | | | | | | | | |
Db 1852 TAAAAAATAAAAAAATAAAAAA 1870

```

Search completed: January 29, 2004, 13:11:27
 Job time : 293.391 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 13:44:42 ; Search time 35.197 Seconds
(without alignments)
3904.448 Million cell updates/sec

Title: US-09-996-617-2
Perfect score: 7534
Sequence: 1 MAGGAWRLACYLEFLKKE.....HLIMELWKGSKKGLPLSS 1429

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5969	79.2	1192	2 T17255	hypothetical prote
2	882	11.7	1111	2 A53000	mater protein [imp
3	465.5	6.2	483	2 S27880	Nasopressin recept
4	361.5	4.8	456	2 S20597	ribonuclease inhib
5	358	4.8	461	2 A31858	ribonuclease-angio
6	354.5	4.7	456	2 A31857	ribonuclease inhib
7	345	4.6	1130	2 A48943	MHC class II trans
8	206	2.7	1075	2 T31668	hypothetical prote
9	202	2.7	1004	2 T31665	hypothetical prote
10	198	2.6	1232	2 A55478	neuronal apoptosis
11	196.5	2.6	1447	2 T42628	neuronal apoptosis
12	180	2.4	533	2 T52063	ran GTPase-activat
13	166.5	2.2	1120	2 JC7765	mitotic spindle as
14	159.5	2.1	506	2 A45841	T-complex-associat
15	155.5	2.1	1199	2 T47442	disease resistance
16	153.5	2.0	1131	2 F96662	hypothetical prote
17	152	2.0	1027	2 T46296	hypothetical prote
18	151	2.0	3511	2 A59295	unconventional myo
19	151	2.0	1221	2 T52347	disease resistance
20	150.5	2.0	1220	2 T48928	disease resistance
21	149	2.0	631	2 C89243	protein F28C1.3 [i
22	149	2.0	631	2 T21471	hypothetical prote
23	148	2.0	1217	2 T52348	disease resistance
24	147.5	2.0	793	2 S54772	mammary gland fact
25	147.5	2.0	1214	2 T47438	disease resistance
26	147	2.0	1017	2 H96663	hypothetical prote
27	147	2.0	1784	2 T10532	gag-pol polyprotei
28	143.5	1.9	4302	2 A38971	polycystic kidney
29	142	1.9	1215	2 T00364	hypothetical prote

adenylate cyclase
RAN GTPase-activat
disease resistance
hypothetical prote
hypothetical large ATP
probable large ATP
probable disease r
transcription acti
adherence factor r
smoothelin - human
p-glycoprotein - S
hypothetical prote
hypothetical prote
disease resistance
hypothetical prote
ATP-dependent prot

30 142 1.9 1839 1 OYBYK
31 139.5 1.9 545 2 T52068
32 139.5 1.9 1189 2 T52346
33 138.5 1.8 1031 2 G96663
34 138 1.8 1010 2 T36383
35 137 1.8 1226 2 T45788
36 136.5 1.8 794 2 G02317
37 136.5 1.8 3335 2 H81702
38 135.5 1.8 1422 2 B71437
39 135 1.8 915 2 T09575
40 135 1.8 1724 2 T18343
41 134 1.8 889 2 H96606
42 133 1.8 1184 2 H71436
43 133 1.8 1301 2 D85188
44 133 1.8 1453 2 G96613
45 132.5 1.8 805 2 AH2731

ALIGNMENTS

RESULT 1
T17255
hypothetical protein DKFZp586O1822.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T17255
R:Kohrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A:Reference number: Z18722
A:Accession: T17255
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1192 <KOE>
A:Cross-references: ENBL:AL117470
A:Experimental source: adult uterus; clone DKFZp586O1822
C:Genetics:
A>Note: DKFZp586O1822.1

Query Match 79.2%; Score 5969; DB 2; Length 1192;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 1147; Conservative 1; Mismatches 0; Indels 44; Gaps 1;

Qy 282 LLLQRPSPDPLVKRSWPDYVEENRGLHIEIRDLFGGLDTQPRIVILQGAAGIGKS 341
Db 1 LLLQRPSPDPLVKRSWPDYVEENRGLHIEIRDLFGGLDTQPRIVILQGAAGIGKS 60
Qy 342 TLARQVKEAWRGQLYGDRFQHVFFSCRELAQSKVSLAEILIGKDGATATPAPIRQILSR 401
Db 61 TLARQVKEAWRGQLYGDRFQHVFFSCRELAQSKVSLAEILIGKDGATATPAPIRQILSR 120
Qy 402 PERLLFILDGVDEPGWVLOEPSSSELCLHWSQOPADALLGSLGKTIILPEASFLITARTT 461
Db 121 PERLLFILDGVDEPGWVLOEPSSSELCLHWSQOPADALLGSLGKTIILPEASFLITARTT 180
Qy 462 ALQNLIPSLQARWVEVLGFSSESSKEYFYFTDERQAIRAFRLVKSNKELWALCLVPW 521
Db 181 ALQNLIPSLQARWVEVLGFSSESSKEYFYFTDERQAIRAFRLVKSNKELWALCLVPW 240
Qy 522 VSWLACTLMQMKRKEKLTITSKTTTTLCHLYLAQALQAOPLGQPLDCLSLAAGIWIQ 581
Db 241 VSWLACTLMQMKRKEKLTITSKTTTTLCHLYLAQALQAOPLGQPLDCLSLAAGIWIQ 300
Qy 582 KKTLESPPDLRKHGLDGAISTFLKMGILQEHPIPLSYSFTHLCFOEFPAAMSVYLEDEK 641
Db 301 KKTLESPPDLRKHGLDGAISTFLKMGILQEHPIPLSYSFTHLCFOEFPAAMSVYLEDEK 360
Qy 642 GRGKHSNCIIDEKTELYAGIHGLFGASTTRFLGLLSDEGEREMENIFHCLRSQGRNLM 701
Db 361 GRGKHSNCIIDEKTELYAGIHGLFGASTTRFLGLLSDEGEREMENIFHCLRSQGRNLM 420
Qy 702 QMWPSLQLLPHSLLESFLYETRNKTLFTQWMAHFEMGMCVETDMELLYCTTCIKPS 761
Db 702 QMWPSLQLLPHSLLESFLYETRNKTLFTQWMAHFEMGMCVETDMELLYCTTCIKPS 761

Db 421 QWVPSLQLLQPHSLBSLHCLYETRNKTLFTQVMAHFEEGMCMVETDMELLVCTFCIKFS 480
Qy 762 RVVKLQLLEGQHRSTWPTWVLPFRWVPVTDAYWQILFVSLKVTNKLKELDLGNSLS 821
Db 481 RVVKLQLLEGQHRSTWPTWVLPFRWVPVTDAYWQILFVSLKVTNKLKELDLGNSLS 540
Qy 822 HSAVSKLCTLRPRCLLETTLRAGCGLTAEDCKOLAFGLRANQTLTDLDFSNVLTDA 881
Db 541 HSAVSKLCTLRPRCLLETTLRAGCGLTAEDCKOLAFGLRANQTLTDLDFSNVLTDA 600
Qy 882 AKHLQRLQPCQKQLQRLQVCGSLTSDCCQDLASVLSASPSLKELDLQNNLDDVGVRL 941
Db 601 AKHLQRLQPCQKQLQRLQVCGSLTSDCCQDLASVLSASPSLKELDLQNNLDDVGVRL 660
Qy 942 LCEGLRHPACKLIRGLDQTLSDENRQELRALEQEKPOLLI FSRKPKSVMTPTGCLDGT 1001
Db 661 LCEGLRHPACKLIRGLDQTLSDENRQELRALEQEKPOLLI FSRKPKSVMTPTGCLDGT 720
Qy 1002 EMSNSTSSLKRORLGSEAAASHVAQANLKL DVSKI FPIAEIAESSPEWVPVELLCVPS 1061
Db 721 EMSNSTSSLKRORLGSEAAASHVAQANLKL DVSKI FPIAEIAESSPEWVPVELLCVPS 780
Qy 1062 PASQGLHTKPLGTDDDFGPTGPVATEVVDVKEKNLYRVHFPVAGSYRWPNTGLCFVME 1121
Db 781 PASQGLHTKPLGTDDDFGPTGPVATEVVDVKEKNLYRVHFPVAGSYRWPNTGLCFVME 840
Qy 1122 AVTVEIEFCWDFGEINPQHSMMVAGPLDIKAEFGAVEAVHLPHFVALOGGHVDTSL 1181
Db 841 AVTVEIEFCWDFGEINPQHSMMVAGPLDIKAEFGAVEAVHLPHFVALOGGHVDTSL 900
Qy 1182 FQMAHFKEGMLLEKPARVELHHIVLENPSFPLGVLLKMIHNALRFIPVTSVLLYHRV 1241
Db 901 FQMAHFKEGMLLEKPARVELHHIVLENPSFPLGVLLKMIHNALRFIPVTSVLLYHRV 960
Qy 1242 HPEEVTFLYLIPSDCSIR - - - - - 1260
Db 961 HPEEVTFLYLIPSDCSIRKAIDLEMKQFVRIHKPPPLTPLYMCRVTVSGSGGML 1020
Qy 1261 ---KELELCYRSGEDOLSEFVGHLSGIRLOVKDKDETLVWEALVKPGDLMPATTL 1317
Db 1021 ILPKELELCYRSGEDOLSEFVGHLSGIRLOVKDKDETLVWEALVKPGDLMPATTL 1080
Qy 1318 IPPARIAPSPDAPOLLHFVQYRQELIARTSVSEVWLDKLGHOVLSOEQYERVLANT 1377
Db 1081 IPPARIAPSPDAPOLLHFVQYRQELIARTSVSEVWLDKLGHOVLSOEQYERVLANT 1140
Qy 1378 RPSQMRKLFSLQSDWRKCKDGLYQALKETHPHLIMELWEKSKGLPLSS 1429
Db 1141 RPSQMRKLFSLQSDWRKCKDGLYQALKETHPHLIMELWEKSKGLPLSS 1192

RESULT 2
A59000
mater protein [imported] - mouse
C:Species: Mus musculus (house mouse)
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C:Accession: A59000
R:Yong, Z.B.; Nelson, L.M.
Endocrinology 140, 3720-3726, 1999
A:Title: A mouse gene encoding an oocyte antigen associated with autoimmune premature ovulation
A:Reference number: A59000; MUID:99360614; PMID:10433232
A:Accession: A59000
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1111 <KUR>
A:Cross-references: GB:AF074018; NID:G5802697; PIDN:AAD51762.1; PID:G5802698
C:Genetics:
A:Genes: Mater

Query Match 11.7%; Score 882; DB 2; Length 1111;
Best Local Similarity 26.1%; Pred. No. 2.4e-49;
Matches 274; Conservative 154; Mismatches 350; Indels 272; Gaps 25;

Qy 162 SSPHESQSPSNAP - - - - - TSTAVLGSWGSPPOPS - - - - - LAPR - - - - - EQAPGQTOWPLD 209
Db 28 TSPENDSKSIQKQOQPEQEQTSESTMG - - - - - PPKDSKAILKARGLEBQESKSTMSPE 83
Qy 210 ETSGLIYITEIREREREKSEKGRPPWAAVGTTPQANTSLQPHHPHWPESVRESLCTWFW 269
Db 84 NVSRAILLDSGSEVEQASERK - - - - - MTSPENDSKSIQKQOQPEQEQTSETLOS - - - - - 133
Qy 270 KNEDFNQFTQLLLQRLPHRPSQDPLVKRSPDVPVEENRGLIE - - - - - 313
Db 134 KEED - - - - - EVTE - - - - - ADQNGGD - LQDYKAHVIAKFTDSVLDHVDSPWK 175
Qy 314 - IRDLFGGLDTQBPRIIVILQCAAGIGKSTLARQVKEARGOLYGRDFQHFVFPSCREL 372
Db 176 LLSDAFKYQKTFQPHITIIHGRPGVGSALARSIVLQMAQKLF - QKMSFVIFSVREI 234
Qy 373 AOSKVVSIAELIGKDGATAPAPIROIILSRPRLILFDGVDEPGWVLQEPSELCLHWSQ 432
Db 235 KWTEKSSLAQIATAKECPDSDWLVTKIMSQPERLLFVIDGLDDMSVLQHDDMTLSRDWD 294
Qy 433 POPADALLGSLGKTLTILPEASFLITARTALQNLIPSLQARWVVEVLGFSSESSKEYFYR 492
Db 295 EOPIYILMYSLLRKALLPQSFLIITRTWTGLEKSMVVSPLYILVBEGLSASRSQVLVE 354
Qy 493 YFTDERQAIRAFRLVKSNNKELWALCLVPWSWLA CTCLMOQMKRKEKLTLSKTTT - - - - - 548
Db 355 NISNESDRIQVPHSLIENHQLFDQCAPSVCSLVCEALQLOKLGKRCRTLPCQTLTGLYA 414
Qy 549 TLUHVL - - - - - AQAQAQPLGPOLRDLCSLAAGSIWQKTLFSPDDTLKRGHGLDGAIIST 603
Db 415 TLVPHQLTKRPSQSALSQEBEQITLVGLCMAAGVMTMRVSFVYDDDLKNTSLKESILA 474
Qy 604 FLKWCILQE - - - - - HPILPSYFHLFCFQEPFAAMSIVLEDEKGRKH - - - - - SNCIIDLEK 655
Db 475 LFHMNILLQVHNESQCVFVSHLSLQDFFAALYYVLEGLEEWNQHFCFIENQRSIMEVYR 534
Qy 656 TLEAYGIHGLGASTRELLGLSDGEGEREMENIFHCRLL - - - - - SOGRNLMQWVPSLQLLQP 713
Db 535 TDDT - - - - - RLGLMK - - - - - RFLGLMKNK ILKTLLEVLPYFVPIVTEQKLOHW - - - - - SLIAQQ 585
Qy 714 - - - - - HSLSLHCLYETRNKTLFTQVMAHFEEGMCMVETDMELLVCTFCIKFSRHWK 766
Db 586 VNGTSPMDTLDAFYCLPESQDEEFVGGALKRFQEWLLINOKMDLKVS SYCLKHQNLKA 645
Qy 767 LQL - - - - - IEGQHRSTWSPVVL - FRWVPVTDAYWQILFVSLKVTNKLKELDLGNSLS 821
Db 646 IRVDIRDLSDVNTLELCFVVTVOETCKPLLMWGNFCSVLGSLRNLKELDLGDSILS 705
Qy 822 HSAVSKL - - - - - 828
Db 706 QRAMKILCLELRNOSCRIOKLTFFKSAEVVSGLKHLLKLLFSNQLKYLNLGNTPMKDDDM 765
Qy 829 - - - - - CKTLRRPRCLLETLR - - - - - 843
Db 766 KLACEALKHPKCSVETLRDSCELTIIGYEMISTILLISTRLKCLSLAKNRVGVKSMISL 825
Qy 844 - - - - - LACGGLTAEDCKDLAFGLRANQTLTDLDFSNVLTDA GAKHLCQR 888
Db 826 GNALSSMCLLQKILDCNCGTTPASCHLLVSALFSNQLTTHLCLSNNSLGTGEGVQOLCOF 885
Qy 889 LRQPSCKLQR - - - - - 898
Db 886 LRNPECALQRLILNHCNIVDDAYGFLANRLANNYKLTHTLSLTMPVPGAMKLLCEALKE 945
Qy 899 - - - - - LQVLSGGLTSDCCQDLASVLSASPSLKELDLQNNLDDVGVRLICEGLRHHPAC 951
Db 946 PTCYLQELVELVDQLTQCCEDLACMITTKHLKSLDLGNALGDKGVITLICEGLKQSSS 1005
Qy 952 KLIRGLDQTLSDENRQELRALEQEKPOL 981
Db 1006 SLRRLGLGACKLTNCCCEALSALSCNPHL 1035

RESULT 3

S27880
Nasopressin receptor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 05-Nov-1999
C:Accession: S27880
R:Herrera, V.L.; Ruiz-Opazo, N.
submitted to the EMBL Data Library, February 1992
A:Description: Characterization of a cDNA encoding an AII and AVP receptor isoform.
A:Reference number: S27880
A:Accession: S27880
A:Molecule type: mRNA
A:Residues: 1-483 <HER>
C:Genetics:
A:Gene: AVP

Query Match 6.2%; Score 465.5; DB 2; Length 483;
Best Local Similarity 30.6%; Pred. No. 1.3e-22;
Matches 161; Conservative 74; Mismatches 186; Indels 105; Gaps 18;
QY 534 MKRKEKLTSTKTTTCLHLYLAQALQ-AQPLQP----QLRDLCSLAAGIWKQKTLFSP 588
Db 1 MELGRDLSTSTKTTSTVYLLFTSMKSGATNGPRVQGEELMLCLAREGILKHQAQFSE 60
QY 589 DDLR--KHGLDGAIIIST-FLKM-----GILQEHPIPLSYFIHLQCFEPFAAMSYVLEDEK 641
Db 61 KDLERLKLKLGSGVQVTFMFLSKKELPGVLE---TVVTYQFIDQSPQEFLLAALSYLLDAB- 116
QY 642 GRGKSNCIIDLEKTLKLVANGIHLFG--ASTRFLGLLSDGERMENIFHC-----RLS 695
Db 117 --GAPGNSAGSVQMLNLSDA--GLRGHLALTTRFLGLLSTERIRDIGNHFGCVPGRYK 172
QY 696 QGRNLQMOWPSQLLQLP-----HSLSLHCLY 723
Db 173 Q--DTLRWV---QGGQPKVATVGAEKKBELKDEABEEEEEEEEELNFGLELLCY 227
QY 724 ETRNKTLFLQVMAHFEMGCVB-----TDMELLVCTFCIKFSRHVKKLQLI-----E 771
Db 228 ETQEDDFVQALSSLPB--WVLERVLTTRMDLEVLVSVYCCQCPDQALRVSCGLVAKE 285
QY 772 GRQHRSTWSTWVLRFRWPVPTDAYWQIILFVLKVTNRNLKELDLSGNSLSHSAVKSLCKT 831
Db 286 KKKKKKSF-----NNRLKGSQSTGKQPAPSLRLPLCEA 318
QY 832 LRPRCLLETLRAGCLTAEDCKDLAFGRANQTLTDLSPNLTLDAGAKHLQORLQ 891
Db 319 MITQQCGSLTILSHCKLPDAVCRDSEALKVAPSLRELGLOQRNLTAEAGRLLSQGLAW 378
QY 892 PSCKLQRLQVSCGLTSDCCQDLASVLSASPSLKELDLQNNLDDVGVRLLCGLRHLPAC 951
Db 379 PKCKVOTLRQMPGL-QEVIHYLVIVLQSPVLTLLDSCQQLPGTVVFLCSALKHPKC 437
QY 952 KLIRLGLDQTTLSDEMQRLEALEQKPKQLLIFSRKPSVMTPEG 997
Db 438 GLKTLSTVELTENPLRELQAVKTLKPDIAIHSKLGTHPQPLKG 483

RESULT 4

S20597
ribonuclease inhibitor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
C:Accession: S20597
R:Kawanomoto, M.; Motojima, K.; Sasaki, M.; Hattori, H.; Goto, S.
Biochim. Biophys. Acta 1129, 335-338, 1992
A:Title: cDNA cloning and sequence of rat ribonuclease inhibitor, and tissue distribution
A:Reference number: S20597; MUID:92162755; PMID:1536887
A:Accession: S20597
A:Molecule type: mRNA
A:Residues: 1-456 <RAW>
C:Cross-references: EMBL:X62528; NID:g57670; PIDN:CAA44388.1; PID:g57671
C:Superfamily: ribonuclease inhibitor; leucine-rich alpha-2-glycoprotein repeat homology

F:280-304/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 4.8%; Score 361.5; DB 2; Length 456;
Best Local Similarity 35.5%; Pred. No. 7.9e-16;
Matches 89; Conservative 32; Mismatches 73; Indels 57; Gaps 2;
QY 792 VTDAYWQILFVLKVTNRNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLETLELACGLTA 851
Db 206 ITSANCKDLCDVAVASKASQLQELDLGSKNLGNTGIAALCSGLLLPSCRLRTLWMLWDCDVTA 265
QY 852 EDCKDIAFGLRANQTLTDLISFNVLTDAGAKHLQORLQROPCKLQRLQVSCGLTS--- 908
Db 266 EGCKDLCTVLRKQSLKSLSLAGNELKDEGAQLLCESLLEPCQCQLSLVWKTCSLTAASC 325
QY 909 -----DC-----CQDL 914
Db 326 PHFCSVLTKNSLFLQMSNPGLDGGVVELCKALGYPTVLRVLWLGDCDVTDSGCSL 385
QY 915 ASVLSASPSLKELDLQNNLDDVGVRLLCGLRHLPACKLIRLGLDQTTLSDEMQRLELRL 974
Db 386 ATVLLANRSLRELDLSNNCMGDNVQLLESILKQPSCLIQQLVLYDIYWTDEVEDQLRAL 445
QY 975 EOEKPOLLIIFS 985
Db 446 EERPSSLRIIS 456

RESULT 5

A31858
ribonuclease-angiogenin inhibitor - human
A:Alternate names: ribonuclease inhibitor, placental
C:Species: Homo sapiens (man)
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 19-May-2000
C:Accession: A31858; S02012; S23933; S48636; T47188
R:Lee, F.S.; Fox, E.A.; Zhou, H.M.; Strydom, D.J.; Vallee, B.L.
Biochemistry 27, 8545-8553, 1988
A:Title: Primary structure of human placental ribonuclease inhibitor.
A:Reference number: A31858; MUID:89118269; PMID:3219362
A:Accession: A31858
A:Molecule type: mRNA
A:Residues: 1-461 <LEE>
A:Cross-references: GB:M22414; NID:g186260; PIDN:AAA59130.1; PID:g307040
R:Schneider, R.; Schneider-Scherzer, E.; Thurnher, M.; Auer, B.; Schweiger, M.
EMBO J. 7, 4151-4156, 1988
A:Title: The primary structure of human ribonuclease/angiogenin inhibitor (RAI) discloses
A:Reference number: S02012; MUID:89210799; PMID:3243277
A:Accession: S02012
A:Molecule type: mRNA
A:Residues: 1-422, 'SE', 425-461 <SCH>
A:Cross-references: EMBL:X13973; NID:g35843; PIDN:CAA32151.1; PID:g35844
A:Note: part of this sequence, including the carboxyl end of the mature protein, was con
R:Crevel-Thieffry, I.; Cotterill, S.; Schuller, E.
Biochim. Biophys. Acta 1122, 107-112, 1992
A:Title: Characterisation of a tryptic peptide from human placental ribonuclease inhibiti
A:Reference number: S23933; MUID:92338217; PMID:1633192
A:Accession: S23933
A:Molecule type: protein
A:Residues: 174-195 <CRE>
R:Nadano, D.; Yasuda, T.; Takeshita, H.; Uchida, K.; Kishi, K.
Arch. Biochem. Biophys. 312, 421-428, 1994
A:Title: Purification and characterization of human brain ribonuclease inhibitor.
A:Reference number: S48636; MUID:94311593; PMID:8037455
A:Accession: S48636
A>Status: preliminary
A:Molecule type: protein
A:Residues: 2-14 <NAD>
R:Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24374
A:Accession: T47188
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-461 <AAA>


```

QY 438 -ALLGSLGKTLPEASFIITARTTALQNLIPSEQA-RWVEVLGFSESSKEYFYRYFT 495
DB 526 RGLLAGLFGKLLRGCTLLLTARPG--RLVQSLSKADALFELSGMEQAQAYVMRYFE 583
QY 496 D-----BROAIRAFVKSNEKWLALCLVPWWSWLAQCLTCLMOQMKREKTLTSTKTTTLC 551
DB 584 SSGMTHQD-RALTLLRDRLPLLSHSHSPTLTCLRAVCQLSEALLEGEADKLPS-TLTGLY 641
QY 552 LHYLAQALQALPGLPQRLDLCSLAAB-GIWQKKT-----FSPDLRKHGLDGAISTFLK 606
DB 642 VGLLGRAALDSPG-ALAEALAKLAWELGRRHQSTLQEDQFSPADVR-----TWAMA 691
QY 607 MGILQHP-----IPLSY-FSIFHCF-QEPPAAMSYYLEDE-----KGRGHSNCI 650
DB 692 KGLVQHPRAASELAPFSLQCFGLGALWALSGBEIKDELPOYLALTPRKKRPYDNWL 751
QY 651 IDEKTEAYGHGLFGASTTRFLGLSDEG-----EREM 686
DB 752 EGVPRFLA-----GLIFQPPARCLGALLGPSAAASVDRKQVLYRLKRLQPGTLRARQL 806
QY 687 ENIFHC-----RLSQGRNLMQWVP-----SIQLLLQPHSLF----- 717
DB 807 LELLHCAHEAEAGIWHQVQELPGRSLFGLTPTLTPDAHVGLKALEAAGQDFSLDRST 866
QY 718 -----SLHCLYETRNKTFQTQVMAHFEEMGMCVETDM-----ELLVCTFCIKF 760
DB 867 GICPSGLSLVGLSCV--TRFRAALSDTVALWESLRQHGETKLLQAAEKEFTIEPPKAKS 924
QY 761 SRHVKKL-QLIEGRQHRSTWSPMVL-----FRWVPVTDAY-WQILFSLVKVTR 808
DB 925 LKDVEDGLKVQVQTRSSSEDAGELPAVRDLKLEFALGPVSGPQAPPKLVRIITAFS 984
QY 809 NLKELD--LSGNSLSHSAVSKLTKLRPRCLLETLRLAGCGLTAEDCKDLAFLGRANQ 865
DB 985 SIQHLDLALSENKIGDEBVSQSATF--PQL-----K 1015
QY 866 TLTELDLSFNVLTDAGAKHLCORLPSCCKLQRLQVSCGLTSDCCQDLASVLSASPSLK 925
DB 1016 SLETLNLSQNNITDLGAYKLAEALPSLAASLLRLSLVNNICIDVGAESLARVLPMVSLR 1075
QY 926 ELDLQNNLDDVGVRLLCGLR 947
DB 1076 VMDVQYNKFTAAGAQAQLAASLR 1097

RESULT 8
T31668
hypoetical protein COSI.5 - sea squirt (Ciona intestinalis)
C:Species: Ciona intestinalis
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 20-Jun-2000
C:Accession: T31668
R:Bird, A.P.; Clark, V.; Jones, S.J.; Leitgeb, S.; Lennard, N.; Tweedie, S.
submitted to the EMBL Data Library, October 1996
A:Reference number: T31668
A:Accession: T31668
A:Molecule type: DNA
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Residues: 1-1075 <BIR>
A:Cross-references: EMBL:Z80904; PIDN:CAB02589.1
C:Genetics:
A:Introns: 48/1; 557/3; 611/3; 685/2; 721/2; 739/1; 779/3; 820/3; 853/2; 914/2; 993/1

Query Match 2.7%; Score 206; DB 2; Length 1075;
Best local Similarity 18.5%; Pred. No. 4.8e-05;
Matches 209; Conservative 152; Mismatches 347; Indels 422; Gaps 49;

QY 301 PDYVEENRGHLIBIRDLFGGLDT-----QEPRIVLQGAAGKST 342
DB 37 PEPAPNVQPLVTDQLFDKALENAQRYTEDESAEYAKWIERHANTVWVGPPGVGKTT 96
QY 343 LARQVKEAMRGQLYGRFOHFYFSCRELAQSKVSVLAELIGKQGT-----TPAP 394
DB 97 LKQWVKQILKHEFLPDT-EYIFFINVKDIDFNKEMTLEFLTTNSRKVKVNYTEESKAL 155

```

RESULT 9
T31665
hypoetical protein COSI.4 - sea squirt (Ciona intestinalis)

```

QY 395 IRQILSRPERLLFILDGVDEPGWVLQEPSSSELCLHWSQPQADALLGSLGKTIILPEASF 454
DB 156 ITFLHNNPVAIF-FDGLDEASTNEFARIPHIKLDGSKPVD-IMKNLNLTLPLPKAI 213
QY 455 LITARTTALQNLIPSEQA-RWVEVLGFSESSKEYFYRYFTDERQAIRAFRLVKSNEKWL 514
DB 214 VVTSTHQWYKLPYRPTSFIEVILGLEEAKNGLTQLCGEKYPAIK--KILDQOPNLA 271
QY 515 ALCLVPWWSWLAQCLTCLMOQMKREKTLTST-----KTTTTLCLHYLAQALQALPGLPQRLD 570
DB 272 HLCYLPINFILVFCLSN-EGSDIKTWTQVLI FMTFRVELSHLKGVPVLDKVGAEWYK 330
QY 571 LCSLAAGIWIQKTLFSPDDLFKHLGDGAISTF-----LKMGILOHPHPLSYS 620
DB 331 LARLAYKGLQORLKFVFKTDFFDVKLADEMVTNFHTYVDKSGGIRMKILEGN--KRSY- 387
QY 621 FHLFCQEPFAAMSIVL----- 637
DB 388 FTHLIWQEFYAAVYLMFLFVSYPREFQLKPIFKDAQWKR VVGFPMGICNPPAYKQLKLVFP 447
QY 638 -----EDEK-----GRG----- 644
DB 448 ATMWKDYBEKKELMVPMMESLWSARGEDLIRRFGLHLEYNDDESSKKEFEDYLPVGLKMD 507
QY 645 --KHNSCIIDLEKTEAY-----GIHGLFGASTTRFLLG-- 676
DB 508 APKHLSEVKDLVYALKSPKPKHLRLDSYETTTTTEVLETLTLLRGVGT-TTITTRFVINNI 566
QY 677 -----LLSDGEREMENI-----FHCRLSQGRNLMQWVPSLQLLQPHSLES 719
DB 567 EMKDSIMELLHLDLAMEBELRFRVTNLSPYMERLSNAIN--QRSNKIQVLVIKH--KL 621
QY 720 HCLYETRNKTFQTQVMAHFEEM--GMCVETDMLLVCTFCIKFSRHVKKQLQIEGRQHR 776
DB 622 HDYDVK--YLAGCLGNISLLYMMGTDISSQ-----CSVLKQAIQQLPSIQ-- 666
QY 777 STWSPTMVFVRWVPVTDAYWQILFSLVKVTRN-----LKDLSGNSLSHSAVSKLCK 830
DB 667 -----VHQLYPDILSTYLNVARNWISNMFVSEVDLVPEALN----- 702
QY 831 TLRRPRCLLETURLAGCGLTAEDCKDLAFLGRANQTLTDLDSFNVLTDAGAKHLCQRLR 890
DB 703 -----LNGNL-----KDSKQQRPRNEELCSQDVSVPAPTQVNYNHYC--N 740
QY 891 QPSCKLQRLQVSCGLTSDCCQDLASVLSASPSLKLDELQNNLDDV----- 937
DB 741 APYHQVRFI-----TQAPILF--MINSNLHPNGLDQVEKIRWLEDESEYLE 784
QY 938 GVRLLCEGLRHPA-----CKLI-----RLGLD 959
DB 785 TPKLQCE--PASLTFKEQVTKIQTHVVLDKETIRCKLLHSDKDGITWEETKTKLEFS 840
QY 960 QTTLS-----DEMQRALRQEKPKQLLIFSRRKPSVMTPEGLD--TGEMS 1004
DB 841 DNFISFQTNFYSWKVIPEFIEKTLAFLYKVKLKSLLYQAGVWVWVKWNLDDVATGE-- 898
QY 1005 NSTSSILKQRLGSERAASHVAQANLKLDSKIFPIAEIAEASSPVPVVELLCVPS-- 1061
DB 899 ---NELKNHFQETWT--IPLSNDLILCLEK-----HENEAEVINI----IPSGKI 941
QY 1062 -PASQ-----GLHTKPLGTDDDQFWGPTGPVATEVVDKEKLYRVHFPVAGSY 1108
DB 942 IFANQLNNSYCNCKFKVDKHS--TNE-----VHLIAKAK-----CGSF 978
QY 1109 RWPNTGLCFWREAVTVEIEFCVWDQ--FLGEINPQHSMMVAGPLLDIK 1155
DB 979 RW-DDDFCF-----PLPITHAVSQDQPTQPSINP--TVIVTGPPTTIR 1019

```

C;Species: Ciona intestinalis
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 20-Jun-2000
C;Accession: T31665
R;Bird, A.P.; Clark, V.; Jones, S.J.; Leitgeb, S.; Lennard, N.; Tweedie, S.
submitted to the EMBL Data Library, October 1996
A;Reference number: T21050
A;Accession: T31665
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1004 <BIR>
A;Cross-references: EMBL:280904; PIDN: CAB02586.1
C;Genetics:
A;Introns: 30/1; 539/3; 593/3; 662/3; 728/3; 769/3; 817/2; 881/2; 960/1

Query Match 2.7%; Score 202; DB 2; Length 1004;
Best Local Similarity 17.7%; Pred. No. 7.9e-05;
Matches 191; Conservative 162; Mismatches 373; Indels 352; Gaps 40;

QY 244 AHTSLQPHHPBPSPVRESLCSLTPWKNEDFNQKFTQLLLQRPSPQDPLVGRSWPDY 303
DB 9 AHAAXKHQHP-----EPAPVNVQPLV----- 30

QY 304 VEENRGLHIEIRDLFGGLDT-----QEPRIVLQGAAGIGKSTLAR 345
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 346 QVKEANGRQLYDRFQHVFPYFSCRELAQSKVSLAELIGKDGTA-----TPAPIRQ 397
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 82 MMVYNQILKHELLPDT-EYFFIYAKDIDFNKEMTLLEFLTTSRVKVNTEESKALITF 140
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 398 ILRSRPERLLFILDGVPQWVLQEPSSSELCHWSQPQADALLGLGKTLIPASFLIT 457
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 141 LHNPNVAIF-PDGLDEASIKELGVYSICKLDEKSPVD-IMKNLFLNALLPKAKIVT 198
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 458 ARTALONILPSLEQARWVEVGFSSSRKEYFYRYFTDERQAIRAFRLVKNKELMALC 517
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 199 STPDENFLNCHQVRPTSIPEVFLGFLBEAKNLTQLCGEKYPAIK--KILDQPNLAHL 256
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 518 LVPWVSWLACTCLM-----QOMRKEKLTLSKTTTTLCHLYLAQALQALQPLQDRDLS 573
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 257 YLPINILVFCLLSNEGSDIKMTQVLIFSMTFRVESH-LKGEVPLDKVGAEMVKLAC 315
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 574 LAAGIWKQKTLFSPDRLKRGHGDGAIISTF-----LKMGIQHPHPIPLSYSTH 623
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 316 LAYKGLQQRKLVPEKTFDDVDKLADENVTFHYVDISSGIRIKLEGN--KRSY-FTH 372
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 624 LCFQEPFAMSVYLEDEKGRGKHSNCIIDLKTLAYGIHGLFGASTTFLGLLSDGE 683
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 373 HIWQEPYAAV-----YLMFLVSYREF 393
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 684 REMENIFHCRLSQGRNLMQWVPSLQLLL---QPHSLSLHCL-----YETRNKTFIT 732
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 394 EOLKTFED-----TQMSVVVFMFGICNPFPAYKQLKIFPATMIKDYBEKKFLES 445
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 733 QVMAHPEEMGMCVETDMELLVCTFCIKRHFGRKQLQIEGROHRSWTSMVVLFRWVPV 792
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 446 MWM---ESLSSAKSEL-----TIRP-----GWLHEYNDDTSKKFDCLPV 484
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 793 TDAYWOILFVLKVTNRLKELDSLGNLSHSAVKSLCKTLRPRCLLETFLRAGCGLTAE 852
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 485 -----GLKMGVPKHLPEVKDLVYALKSFTKPKHLRLRSNWTTTTEVLETLRGIHGTIT 539
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 853 DCKDLAFGLRANQTLTELDLSFNVLTDAGAKHLQBLRQPSCKLQBLQVLSGCLTSDCCQ 912
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 540 ITRFVNNTEMKDSLMELEL---LHLDA-----MEKLRFPDDVTNLSYME 581
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 913 DLASVLSASPSKELDL---QQNNLDVGVLLCEGLRHPACKLIRGLDQDTTLLSDMRQ 969
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 582 SILSNAINQRNKIQLDLWHQQLNDVVLACGLG-----NISLANSHTYIISD--- 632
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 970 ELRALQEQKPOLLIIFSRKPSVMTPTTEGLDTGEMSNSTSLKQRQLGSRASASHVAQNL 1029
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 633 QCRVLKQATEQL-----PSI----- 647
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 1030 KLLDYSKIFPIAETAESESPVVPVELLCVSPASQGLHTKPLCTGDDDFWGTGPVATE 1089
DB 648 ---QVHQLY-----PDILST-YLNVVRPIIRFDFT----- 674

QY 1090 VVDKEKNLYRVHFPVAGSYRWPNTGLCFVMBEAVTVIEFCVWDQFLGEINPQHSHMWAG 1149
DB 675 -----SVYFVH-----DQPFSSSKCWIIGRGGKLEVGCC----- 704

QY 1150 PLLDIKAEFGAVE---AVHLPHFVALQGHVDTSLFOWAHFKEEGMLLEKPARVELH-HI 1205
DB 705 ---ELVVPFGALEKNDVEIKLTASLSLESEFLEPTLQC---ELASLTLLKKQVTKLQTHV 758

QY 1206 VLENPSFSPGLGVLLKMIHNLRFIPVTSVLLLYHRVHPEEVEFHL---YLIPSD-CST 1259
DB 759 VLDKET-----IRCKVTLVYTRV---TTVHWGKGLNHLNHTDICS 796

RESULT 10
A55478
neural apoptosis inhibitory protein - human
N:Alternate names: NAIP
C:Species: Homo sapiens (man)
C;Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 02-Feb-2001
C;Accession: A55478
R;Roy, N.; Mahadevan, M.S.; McLean, M.; Shutler, G.; Yaraghi, Z.; Farahani, R.; Baird, S.
Cell 80, 167-178, 1995
A;Title: The gene for neuronal apoptosis inhibitory protein is partially deleted in indi
A;Reference number: A55478; MUID:95112344; PMID:7813013
A;Accession: A55478
A;Molecule type: mRNA
A;Residues: 1-1232 <ROY>
A;Cross-references: GB:U19251
C;Genetics:
A;Gene: GDB: SMA@; SMA
A;Cross-references: GDB:120378; OMIM:600354; OMIM:253300
A;Map position: 5q12.2-5q13
C;Keywords: apoptosis; ATP; glycoprotein; nucleotide binding; P-loop; transmembrane prote
F;94-110/Domain: transmembrane #status predicted <TM1>
F;470-477/Region: nucleotide-binding motif A (P-loop)
F;479-496/Domain: transmembrane #status predicted <TM2>
F;476/Binding site: ATP (lys) #status predicted
F;618,632,823,923,1035/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 2.6%; Score 198; DB 2; Length 1232;
Best Local Similarity 20.8%; Pred. No. 0.0002;
Matches 158; Conservative 108; Mismatches 282; Indels 212; Gaps 34;

QY 292 QDPLVKRSPDYVEENRGLHIEIRDLFGPLDTPQPRIVILQGAAGIGKSTLARQVKEAW 351
DB 450 QEPLV-----LPEVFG-NLNS-----VNCVEGEGAGSKTVLLKIAFLW 487

QY 352 GRGQL-YGDRFQHVFPYFSCRELAQSK---VVSALAEIGKDGTTATPAPIRQLSR-PERLL 406
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 488 ASGCCPLNRRQLFYFLSLSTRPDEGLASITCDLLEKEGSEVTMCMRNIIQQLKNQVL 547
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 407 FILDGVDEPGWVLPQSSSELCHWSQPQADALLGLGKTLIPASFLITARTALQNL 466
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 548 FLDDY-----KEIC---SIPO---VIGKLTQKNHLSRTCLLIARVTRNARDI 589
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 467 IPSLEQARWVEVGFSSSRKEYFYRYFTDERQAIRAPRL-VKSNKELMALCLVPWVSWL 525
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 590 RRYLETI---LEIQAPFFYNTVCIIRKLFSHNTRURKFMVYFGKNSLOKTKTLPFVA 647
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 526 ACTCLMQQ-----MKRKEKLTLSKTTTTLCHLYLAQALQALQPLQDRDLS 573
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 648 ICAHWFOYFPDPSPDVAVFYSYMERLSRNKAT-----AELKATVSSCGE 694
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 574 LAAGIWKQKTLFSPDRLKRGHGDGAIISTFLKMGILQEHPIPLSYSTHLCFQEPFAM 633
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

QY 695 LALKGFFSCCFEFDNDDLAAGVDEDEDLTMCLSKFTAQRLRPFRFLSPAFQBLFAGM 754
DB -----TIDQLFDKALENAQCYTEDQSAEYAKWYIERHANTVMVWVGHGGKTTLLK 81

Db 257 EAAKVAELIPSTEKLVLFHNNMTGDEGAFAIAEWKRSPALEDFRC---SSTRVSGE 313
 Qy 733 QVMAHFEEMCMCVETDMELLVCTFCIKFSRHVKKQLQIEGRQHRSTWSTMVVFWVPV 792
 Db 314 GGVALAEALGACT-----HLKKDL----- 333
 Qy 793 TDAYWQI-----LFSVLKVTNRLKELDLSGNSLSHSAVSKLCTLRPRCLLETTLRLAGC 847
 Db 334 RDNMFGEAGVALSKVIPFADLTIELVSYLNLEDDGAALANALAKESAPSLETLDAGN 393
 Qy 848 GITAEDCKDLAFGLRANQTLTDLDSFNVLTDAGAKHLQRLQPSCKLQRLQVSCGLT 907
 Db 394 DITAKATVSAEICSSKQFLTKLNLSENELKDEGA-----GL- 430
 Qy 908 SDCCQDLASVLSNPSKELDLQONLDDVGVRLICEG-LRHPACKLIRLGLDQTLTSLDE 966
 Db 431 -----ISKALEGRGOLSEVDLSTNLITWSGAKLAEAVVQKPGFKL--LNINANFISDE 482
 Qy 967 MRQELRALQEKPOLLIFFSRKPSVMTPTTEGLDTGEMSNSTS 1008
 Db 483 GIDELKDIFFNSPDMU-----GFLDNDPEGEDVDDEAEDDS 519

RESULT 13
 JC7765
 mitotic spindle associated protein, MAP126 - human
 C:Species: Homo sapiens (man)
 C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
 R:Chang, M.S.; Huang, C.J.; Chen, M.L.; Chen, S.T.; Fan, C.C.; Chu, J.M.; Lin, W.C.; Yan
 Biochem. Biophys. Res. Commun. 287, 116-121, 2001
 A:Title: Cloning and characterization of hMAP126, a new member of mitotic spindle-associated
 A:Reference number: JC7765; PMID:11549262
 A:Contents: Testis
 A:Accession: JC7765
 A:Molecule type: mRNA
 A:Residues: 1-1120 <CHA>
 A:Cross-references: GB:AF345347
 C:Comment: This protein, as a nuclear protein and a new member of kinesin family, which
 and dynamic regulation of mitotic spindles and may provide an outward force to push the
 C:Genetics:
 A:Gene: hmap126
 A:Map position: 17q11-12

Query Match 2.2%; Score 166.5; DB 2; Length 1120;
 Best Local Similarity 18.8%; Pred. No. 0.02;
 Matches 217; Conservative 138; Mismatches 391; Indels 411; Gaps 50;

Qy 148 WR--EISASLLYQALPSSPDHSPSPQESPNAPTSTAVLGSWGSPQPSLAPR-----EQE 200
 Db 2 WRVKLSLSLSPSPQTKFSMRTPLELTLPQALTTSGK-RSPACSSLTPLSLKLGLOE 60
 Qy 201 AFGTOWPLD-----ETSGIYYTE-----IREREREKSKGRPPMAVVGTPPQAHTSLQ 249
 Db 61 GSNSSPVDVFNKRTDLSSEHFSHSSKWLCTQCHESDEQPLDPIQISSTKTSKTSBEAVD 120
 Qy 250 PHHPHPSPVRESLCSLTWPKWNEFNQKFTQLLQRPHPRSQDPLVKRSDPYVEENRG 309
 Db 121 PLG-----NYMVKTVLVPSPGLGQODM----- 143
 Qy 310 HLIEIRDLFGPLDTQ-EPRIVITLQAGAGIGKSTLARQVKEAWGRQQLYGRFQHVFPYS 368
 Db 144 -----IFEARLDTWAEVNSISLNGP--LRDLDLVREEV-----APCMGDRF----- 182
 Qy 369 CRELAQSKVSVLAELIKGDKGTATPAPIRQLSRPERLLFLDGVDPGVQLQPSSELC- 427
 Db 183 -----SEVAAVSE-----KPIFO-----ESPShLLESPNPPCS 211
 Qy 428 ---LHWSQOPADALLGSLGKLTILPEASFLITARTTAL--ONLIPSLQA-----RWREV 478
 Db 212 EQLHCKSKES-----LSRTEAVREDLVPSSENAFLPSSVLW----- 247
 Qy 479 LGFSESSRKEYFYRYFTDERQAI-----RAPR-----LVKGNKELWALC 517

Db 248 LSPSTALAAADFRVNHVDPEEIEIVHGAMEERMPPTHPKESETEDOALVSSVEDILSTC 307
 Qy 518 LVP-----WVSHLACTCLMQMKRKEKLTLSKTTT 549
 Db 308 LTPNLIVMESQEPAGPAVEDVGRILGSDTESMWSPLA-----WLEKGVNTS 353
 Qy 550 LCLHLYLAQALQAQPLGPQRLDLC-----SLAAEGIMQKTLFSPD-----DLRKHGL 596
 Db 354 VMLNLRSLGSL--LPSMLRDAAGTTTFTSCSVGTW-----FTPSAQEKSTNTSQTGL 405
 Qy 597 DGAIISTFKMGILQEHPIPLSYFSIH-----LCFQEFFAAMSIVLEDEKGR----- 643
 Db 406 VGTKHSTSETQLCGRPDLTALSRLHDLNLLSSLVVEF---LSRQLRDWKSQALVP 462
 Qy 644 -----GKISNCIIDLKTLKAYGHLFGASTTFLGLLSDGEGRENERENIFH 691
 Db 463 HPETQDSSTQTDTSHGITNKLOHLKESH-----EMGQ----- 495
 Qy 692 CRLSGRNLMO-WVPSLQLLQPHSLHSLHCLYETRNKTLTQ----- 733
 Db 496 -ALQOARNVMSWVLSKELI---SLHLHLHLEEDKTTVNOSSRRRAETLVCCCFDLK 551
 Qy 734 -----VMAHFEEMCMCVETDMELLVCTFCIKFSRHVKGLQ--LIEGRQHR 777
 Db 552 KLRAKLQSLKAEREERHREEMALRGDAABIVLEAFCAHASQRIQLEODLASMRFRG 611
 Qy 778 TWSPTMVVFWVPVTDAYWQILFSLVKNLKLDELDSGNSLSHSAVSKLCTLRPRC 837
 Db 612 LLKDAQTLQ---VGLHAKQBELVQQTSLTSLQO--DWRSMLQDYTTWTAL---LSRSRQ 664
 Qy 838 LLETLRLAGCGLTAECDCKDLAFGLRANQTLTDLDSFNVLTDAG--AKHLQRLQPSCK 895
 Db 665 LTKLT-----VKSQALQERDVAIEEKQEVSRVLSQVSNQLECKGQ 707
 Qy 896 LQRLQVSCGLTSDCCDLASVLSASPSLKDLDQNN--LDDVGVV--LLCE----- 944
 Db 708 TEQLELENIRLATDLRAQLANMDSQLKELQSHQTHCAQDLAMKDELLCOLTQSNBEQ 767
 Qy 945 -----GLRHPACKLIRLGLDQTLTSLDQMRQLRALE-----QE 977
 Db 768 AAQCVKEEMALKHMQAELOQ---QQAVALAKEVRDLKETFEADQENQVAHLELQVQECOL 824
 Qy 978 KPQLLIFFSR-----KPSVMTPTTEGLDTGEMSNSTSLLKQRLGSRASASHVAQANLK 1030
 Db 825 KTTLEVLERSLQENLADTVENLTAKLASTIADNOEQDLEKTROYSKLGLLTQ----- 880
 Qy 1031 LLDVSKIFPIAIEAESPEVVPVELLCVPSASQGLHTKPLGTDGDDFWGPTGPVATEV 1090
 Db 881 -LQSLTLFLQTKLKEKTEQETLLLTACPT-----QEHPLPNDRTF---LGSILTAV 929
 Qy 1091 VDKEKNLVRVHPVAGS 1107
 Db 930 ADEEPESTPV--PLIGS 944

RESULT 14
 A45841
 T-complex-associated-testes-expressed-1 protein - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 05-Nov-1999
 R:Accession: A45841
 R:Sarvetnick, N.; Tsai, J.Y.; Fox, H.; Pilder, S.H.; Silver, L.
 Immunogenetics 31, 283-284, 1990
 A:Title: A mouse chromosome 17 gene encodes a testes-specific transcript with unusual pr
 A:Reference number: A45841
 A:Accession: A45841
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-506 <SAR>
 A:Cross-references: GB:M28821; NID:g201909; PIDN:AAA40406.1; PID:g201910

Query Match 2.1%; Score 159.5; DB 2; Length 506;

Db 918 SFPEI 922

Search completed: January 29, 2004, 13:49:45
Job time : 40.197 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:40:41 ; Search time 21.9982 Seconds
(without alignments)
3054.855 Million cell updates/sec

Title: US-09-996-617-2
Perfect score: 7534
Sequence: 1 MAGGAWRLACYLEFLKKE.....HLIMELWEKSKGLPLSS 1429

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	7502	99.6	1473	1 NALI_HUMAN	Q9C000 homo sapien
2	1382.5	18.4	1062	1 PYAT_HUMAN	P59046 homo sapien
3	1279	17.0	1034	1 CISI_HUMAN	Q9P200 homo sapien
4	1254.5	16.7	1033	1 CISI_MOUSE	Q8R4b8 mus musculus
5	1037	13.8	994	1 NAL4_HUMAN	Q96m22 homo sapien
6	983	13.0	1200	1 MATE_HUMAN	P59047 homo sapien
7	948	12.6	1062	1 NAL2_HUMAN	Q9nx02 homo sapien
8	908.5	12.1	980	1 PYA3_HUMAN	Q8wx94 homo sapien
9	901.5	12.0	892	1 PYA5_HUMAN	P59044 homo sapien
10	895	11.9	854	1 PYA5_RAT	Q63035 rattus norv
11	882	11.7	1111	1 MATE_MOUSE	Q9rlm5 mus musculus
12	868	11.5	843	1 PYA5_MOUSE	Q91w82 mus musculus
13	711	9.4	1033	1 PYA6_HUMAN	P59045 homo sapien
14	668	8.9	431	1 CAR8_HUMAN	Q9Y2g2 homo sapien
15	393.5	5.2	1040	1 CARF_HUMAN	Q9hc29 homo sapien
16	392	5.2	1020	1 CARF_MOUSE	Q8k3z0 mus musculus
17	388.5	5.2	953	1 CAR4_HUMAN	Q9Y239 homo sapien
18	361.5	4.8	456	1 RINI_RAT	P29315 rattus norv
19	358	4.8	460	1 RINI_HUMAN	P13489 homo sapien
20	354.5	4.7	456	1 RINI_PIG	P10775 sus scrofa
21	351	4.7	953	1 CAR4_MOUSE	Q8bbh0 mus musculus
22	345	4.6	1130	1 C2TA_HUMAN	P33076 homo sapien
23	318	4.2	1155	1 C2TA_MOUSE	P79621 mus musculus
24	253.5	3.4	193	1 ASC_MOUSE	Q9epb4 mus musculus
25	250.5	3.3	195	1 ASC_HUMAN	Q9ulz3 homo sapien
26	215	2.9	1024	1 CARC_HUMAN	Q9npp4 homo sapien
27	212.5	2.8	1403	1 BIRE_MOUSE	Q9r016 mus musculus
28	200	2.7	1403	1 BIRE_MOUSE	Q9jib6 mus musculus
29	199	2.6	1402	1 BIRC_MOUSE	Q9jib3 mus musculus
30	198	2.6	1403	1 BIRI_HUMAN	Q13075 homo sapien
31	196.5	2.6	1447	1 BIRI_MOUSE	Q9quk4 mus musculus
32	196	2.6	1403	1 BIRA_MOUSE	Q9gw5 mus musculus
33	161	2.1	483	1 YAX2_HUMAN	Q95522 homo sapien

RESULT 1
NALI_HUMAN
ID NALI_HUMAN STANDARD; PRT; 1473 AA.
AC Q9C000; Q9BZZ8; Q9BZZ9; Q9HAV8; Q9UFT4; Q9Y2E0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE NACHT-, LRR- and PYD-containing protein 2 (Death effector filament-forming ced-4-like apoptosis protein) (Nucleotide-binding domain and caspase recruitment domain) (Caspase recruitment domain protein 7).
DE NALP1 OR DEFCAP OR NAC OR CARD7 OR KIAA0926.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=21169419; PubMed=11270363;
RA Bertin J., DiStefano P.S.;
RT "The PYRIN domain: a novel motif found in apoptosis and inflammation proteins";
RL Cell Death Differ. 7:1273-1274(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=21148093; PubMed=11250163;
RA Martinon F., Hofmann K., Tschopp J.;
RT "The pyrin domain: a possible member of the death domain-fold family implicated in apoptosis and inflammation.";
RL Curr. Biol. 11:R118-R120(2001).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX TISSUE=Erythrocyte; PubMed=11076957;
RA Hlaing T., Guo R.-F., Dilley K.A., Loussia J.M., Morrish T.A., Shi M.M., Vincenz C., Ward P.A.;
RT "Molecular cloning and characterization of DEFCAP-L and -S, two isoforms of a novel member of the mammalian Ced-4 family of apoptosis proteins";
RL J. Biol. Chem. 276:9230-9238(2001).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3 AND 4), AND PROTEIN INTERACTION. TISSUE=T-cell;
RX MEDLINE=21153743; PubMed=1113115;
RA Chu Z.-L., Pio F., Xie Z., Welsh K., Krajewska M., Krajewski S., Godzik A., Reed J.C.;
RT "A novel enhancer of the Apaf1 apoptosome involved in cytochrome c-dependent caspase activation and apoptosis.";
RL J. Biol. Chem. 276:9239-9245(2001).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX TISSUE=Brain;
MEDLINE=99246063; PubMed=10231032;
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIII. The complete sequences of 100 new cDNA clones from brain which code

ALIGNMENTS

34	158	2.1	5065	1	EPPL_HUMAN	P58107 homo sapien
35	152	2.0	3511	1	MY15_MOUSE	Q9qz24 mus musculus
36	147.5	2.0	793	1	ST5A_MOUSE	P42230 mus musculus
37	146.5	1.9	1524	1	Y133_HUMAN	Q14146 homo sapien
38	145.5	1.9	793	1	ST5A_RAT	Q62771 rattus norv
39	144	1.9	2300	1	CYAA_NEUCR	Q01631 neurospora
40	142.5	1.9	4303	1	PKD1_HUMAN	P98161 homo sapien
41	142	1.9	1426	1	NPH4_HUMAN	P75161 homo sapien
42	142	1.9	1839	1	CYAA_SACKL	P23466 saccharomyc
43	141.5	1.9	1021	1	PKR_DAUCA	Q81pb4 daucus caro
44	140	1.9	794	1	ST5A_BOVIN	Q95115 bos taurus
45	138.5	1.8	794	1	ST5A_HUMAN	P42229 homo sapien

RT for large proteins, in vitro.";

RL DNA Res. 6:63-70(1999).

[6]

RP SEQUENCE OF 282-1473 FROM N.A. (ISOFORM 1).

RC TISSUE=Uterus;

RA Koehler K., Beyer A., Mewes H.-W., Gassenhuber J., Wiemann S.;

RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

CC -! FUNCTION: Able to form cytoplasmic structures termed death

CC effector filaments. Enhances APA1 and cytochrome c-dependent

CC activation of pro-caspase-9 and consecutive apoptosis. Seems to

CC bind ATP.

CC -! SUBUNIT: Interacts strongly with caspase 2, weakly with caspase 9

CC and with APA1 in a cytochrome c-inducible way leading to the

CC formation of an apoptosome. This interaction may be ATP-dependent.

CC -! SUBCELLULAR LOCATION: Cytoplasmic.

CC -! ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=4;

CC Name=1; Synonyms=NAC beta, DEFCAP-L;

CC IsoId=Q9C000-1; Sequence=Displayed;

CC Name=2; Synonyms=NAC alpha, DEFCAP-S;

CC IsoId=Q9C000-2; Sequence=VSP_004327;

CC Name=3; Synonyms=NAC gamma;

CC IsoId=Q9C000-3; Sequence=VSP_004326, VSP_004327;

CC Name=4; Synonyms=NAC delta;

CC IsoId=Q9C000-4; Sequence=VSP_004326;

CC -! TISSUE SPECIFICITY: Widely expressed. Isoforms 1 and 2 are

CC expressed in peripheral blood leukocytes, chronic myelogenous

CC leukemia cell line K-562, followed by thymus, spleen and heart.

CC Also detected in lung, placenta, small intestine, colon, kidney,

CC liver and muscle.

CC -! SIMILARITY: Contains 1 DAPIN domain.

CC -! SIMILARITY: Contains 1 NACHT domain.

CC -! SIMILARITY: Contains 1 CARD domain.

CC -! SIMILARITY: Contains 7 leucine-rich (LRR) repeats.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement. See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch.

CC -----

DR EMBL; AF298548; AAG15254.1; -

DR EMBL; AF310105; AAG30288.1; -

DR EMBL; AF229059; AAK00748.1; -

DR EMBL; AF229060; AAK00749.1; -

DR EMBL; AF229061; AAK00750.1; -

DR EMBL; AF229062; AAK00751.1; -

DR EMBL; AB023143; BAA76770.1; -

DR EMBL; AL117470; CAB55945.1; -

DR PIR; T17255; T17255.

DR HSSP; P13489; 1A4Y.

DR MIM; 606636; -

DR GO; GO:0005622; C:intracellular; IC.

DR GO; GO:0016506; F:apoptosis activator activity; NAS.

DR GO; GO:0008656; F:caspase activator activity; NAS.

DR GO; GO:0013899; F:enzyme binding activity; IPI.

DR GO; GO:0006919; P:caspase activation; NAS.

DR GO; GO:0006917; P:induction of apoptosis; NAS.

DR InterPro; IPR001315; CARD.

DR InterPro; IPR000767; Disease_resist.

DR InterPro; IPR001611; LRR.

DR InterPro; IPR007091; LRR_RNinh.

DR InterPro; IPR007111; NACHT_NTPase.

DR InterPro; IPR004020; PAAD_DAPIN_dom.

DR Pfam; PF00560; LRR; 2.

DR Pfam; PF02758; PAAD_DAPIN; 1.

DR PRINTS; PR00364; DISEASERSIST.

DR PROSITE; PS50209; CARD; 1.

DR PROSITE; PS50824; DAPIN; 1.

DR PROSITE; PS50837; NACHT; 1.

APoptosis; ATP-binding; Leucine-rich repeat; Repeat;

Alternative splicing.	1	92
FT DOMAIN	328	637
FT DOMAIN	704	725
LRR 1.	807	830
LRR 2.	864	887
LRR 3.	921	944
LRR 4.	950	973
LRR 5.	1199	1215
LRR 6.	1216	1236
LRR 7.	1374	1463
CARD	334	341
NP_BIND	958	987
FT VARSPLIC	1262	1305
FT MUTAGEN	340	340
FT MUTAGEN	340	340
FT CONFLICT	155	155
FT CONFLICT	246	246
FT CONFLICT	782	782
FT CONFLICT	878	878
FT CONFLICT	995	995
FT CONFLICT	1119	1119
FT CONFLICT	1184	1184
FT CONFLICT	1241	1241
FT CONFLICT	1366	1366
FT CONFLICT	1473	1473
SQ SEQUENCE	1473	1473

Query Match 99.6%; Score 7502; DB 1; Length 1473;

Best Local Similarity 97.0%; Pred. No. 0;

Matches 1429; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

QY 1 MAGGAWGLACVLFKKEELKEFQLLLANKAHSSSSGETPAQPKETSGMEVASYLVAQ 60

DB 1 MAGGAWGLACVLFKKEELKEFQLLLANKAHSSSSGETPAQPKETSGMEVASYLVAQ 60

QY 61 YGEQAWDLALHTWEQMLRSILCAQAQEGAGHSFPYSPSEPHLGSPQSTSTAVLMPW 120

DB 61 YGEQAWDLALHTWEQMLRSILCAQAQEGAGHSFPYSPSEPHLGSPQSTSTAVLMPW 120

QY 121 IHELPAQTGQSGRRVRLQLPDTSGRRWRREISASLLYQALPSSPDHSPSESNAPTST 180

DB 121 IHELPAQTGQSGRRVRLQLPDTSGRRWRREISASLLYQALPSSPDHSPSESNAPTST 180

QY 181 AVLGWSGSPPOPSLAPREQEAPGTQWPLDETSGIYYTIREREREKSKGPPMAAVVGT 240

DB 181 AVLGWSGSPPOPSLAPREQEAPGTQWPLDETSGIYYTIREREREKSKGPPMAAVVGT 240

QY 241 PPQAHTSLQPHHPWEPSPVRESLCSCTWPWKNEDEFNQKFTQLLLQRPSPRQDPLVKRSW 300

DB 241 PPQAHTSLQPHHPWEPSPVRESLCSCTWPWKNEDEFNQKFTQLLLQRPSPRQDPLVKRSW 300

QY 301 PDYVEENRGLHIEIRDLFGPGLDTQEPRIVLQGAAGIKGTSLARQVKEAWRGOLYGD 360

DB 301 PDYVEENRGLHIEIRDLFGPGLDTQEPRIVLQGAAGIKGTSLARQVKEAWRGOLYGD 360

QY 361 FQHVYFSCRLAOKSVSLAELIGKDGATATPATRIQLTSRPERLLFLDGDVDEPGWLQ 420

DB 361 FQHVYFSCRLAOKSVSLAELIGKDGATATPATRIQLTSRPERLLFLDGDVDEPGWLQ 420

QY 421 EPSSELCLHWSQOPADALLGLGKTLTPASFLITARTTALQNLIPSLQEARWEVLG 480

DB 421 EPSSELCLHWSQOPADALLGLGKTLTPASFLITARTTALQNLIPSLQEARWEVLG 480

QY 481 FSESRKEYFYRYFTDERQAIAPRLVKSNKELWALCLVPVWSMLACTCLMQMKRKEKL 540

DB 481 FSESRKEYFYRYFTDERQAIAPRLVKSNKELWALCLVPVWSMLACTCLMQMKRKEKL 540

QY 541 TLTSKTTTTLCHVLAQALQAPLGPQLRDLCSLAAGIWOKKTLFSPDDLRKHLGDGI 600

DB 541 TLTSKTTTTLCHVLAQALQAPLGPQLRDLCSLAAGIWOKKTLFSPDDLRKHLGDGI 600

QY 601 ISTFLKMGILQHEPIPLSYFTHLCFOEFAAMSVYLEDEKGRGKSHNCIIDLKTLBAY 660
 DB 601 ISTFLKMGILQHEPIPLSYFTHLCFOEFAAMSVYLEDEKGRGKSHNCIIDLKTLBAY 660
 QY 661 GHGLFGASTTFLGLISDEGEREMENIFHCLSGRNLQMWPSLOLLLOPHSLESILH 720
 DB 661 GHGLFGASTTFLGLISDEGEREMENIFHCLSGRNLQMWPSLOLLLOPHSLESILH 720
 QY 721 CLYETRNKTFQLQVMAHFEEMGCVETDMLLVCTFCIKFSRHVKKLQIEGQRHRSWTS 780
 DB 721 CLYETRNKTFQLQVMAHFEEMGCVETDMLLVCTFCIKFSRHVKKLQIEGQRHRSWTS 780
 QY 781 PTMVVLFWRVPVTDAYQWILFVSVLKVTRNLKELDLSGNSLSHSAVSKLCTLRPRCLLE 840
 DB 781 PTMVVLFWRVPVTDAYQWILFVSVLKVTRNLKELDLSGNSLSHSAVSKLCTLRPRCLLE 840
 QY 841 TLRLAGCGLTAECDKDLAFGLRANQTLTDLTDFSVNLTLDAGAKHLQRLRQPSCKLQRLQ 900
 DB 841 TLRLAGCGLTAECDKDLAFGLRANQTLTDLTDFSVNLTLDAGAKHLQRLRQPSCKLQRLQ 900
 QY 901 LVSCGLTSCCCODLASVLSASPSLKELDLQNNLDVGVRLLCCEGLRHPACKLIRLGLDQ 960
 DB 901 LVSCGLTSCCCODLASVLSASPSLKELDLQNNLDVGVRLLCCEGLRHPACKLIRLGLDQ 960
 QY 961 TTLSDEMROELRALBOEKPOLLIIFSRKPSVMTPTGLDTGEMSNSTSLKQRLGSERA 1020
 DB 961 TTLSDEMROELRALBOEKPOLLIIFSRKPSVMTPTGLDTGEMSNSTSLKQRLGSERA 1020
 QY 1021 ASHVAQANIKLDVSKIFPIAIEAESSPEVVPVELLCVPSASQDGLTKPLGTDDDFW 1080
 DB 1021 ASHVAQANIKLDVSKIFPIAIEAESSPEVVPVELLCVPSASQDGLTKPLGTDDDFW 1080
 QY 1081 GPTGPVATEVDKKNLYRHPFVAGSYRWPNTGLCFVNRBAVTVEIEFCVMDQFLGEIN 1140
 DB 1081 GPTGPVATEVDKKNLYRHPFVAGSYRWPNTGLCFVNRBAVTVEIEFCVMDQFLGEIN 1140
 QY 1141 PQHSMVAGPLDIIKAEPGAFAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
 DB 1141 PQHSMVAGPLDIIKAEPGAFAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
 QY 1201 ELHHIVLENPSPPLGVLLKMHINLALRFPVTSVLLYRHPVEEVTFLYIIPSDCSIR 1260
 DB 1201 ELHHIVLENPSPPLGVLLKMHINLALRFPVTSVLLYRHPVEEVTFLYIIPSDCSIR 1260
 QY 1261 -----KELELCYRSPGSDQLF 1276
 DB 1261 KAIDDLKMFQFVRIHKPPPLPLYMGCRYTVSGSGSGLMEILPKELELCYRSPGSDQLF 1320
 QY 1277 SEFYVGHLSGIRLOVKDKDETLVWEALVKPGDLMPATTLIPPARIAVPSPLDAPQLLH 1336
 DB 1321 SEFYVGHLSGIRLOVKDKDETLVWEALVKPGDLMPATTLIPPARIAVPSPLDAPQLLH 1380
 QY 1337 FVDQYREQLIARTSVTEVVDKHLGQVLSQEQYERVAENRTPSQMRKLFSLQSQWDRKC 1396
 DB 1381 FVDQYREQLIARTSVTEVVDKHLGQVLSQEQYERVAENRTPSQMRKLFSLQSQWDRKC 1440
 QY 1397 KDGLYQALKEHPLHLMELWEKSKKGLPLSS 1429
 DB 1441 KDGLYQALKEHPLHLMELWEKSKKGLPLSS 1473

RESULT 2

PYA7_HUMAN

ID PYA7_HUMAN STANDARD; PRT; 1062 AA.

AC P59046;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE PYRIN-containing APAF1-like protein 7 (Monarch-1).

GN PYPAF7.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI_TaxID=9606;
 (1)
 SEQUENCE FROM N.A. (ISOFORM 1).
 MEDLINE=22162427; PubMed=12019269;
 Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,
 Lora J.M., Geddes B.J., Briskin M., DiStefano P.S., Bertin J.,
 "PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates
 activation of NF-kappa B and caspase-1-dependent cytokine
 processing.";
 J. Biol. Chem. 277:29874-29880 (2002).
 (2)
 SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
 TISSUE=Lymphoma;
 Williams K.L., Linhoff M.W., Harton J.A., Ting J.P.Y.;
 Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 (3)
 SEQUENCE FROM N.A. (ISOFORM 1).
 TISSUE=Leukocyte;
 MEDLINE=22388257; PubMed=12477932;
 Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 Hopkins R.F., Jordan H., Moore T., Max J., Wang J., Hsieh F.,
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 Villalón D.M., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 (4)
 FUNCTION: May mediate activation of CASP1 via ASC and promote
 activation of NF-kappa-B via IKK.
 (5)
 SUBUNIT: Binds to ASC with its DAPIN domain.
 (6)
 SUBCELLULAR LOCATION: Cytoplasmic.
 (7)
 ALTERNATIVE PRODUCTS:
 Event=Alternative splicing; Named isoforms=3;
 Name=1; Synonyms=1;
 IsoId=P59046-1; Sequences=Displayed;
 Name=2; Synonyms=11;
 IsoId=P59046-2; Sequences=VSP_005524;
 Name=3; Synonyms=111;
 IsoId=P59046-3; Sequences=VSP_005523;
 (8)
 TISSUE SPECIFICITY: Detected only in peripheral blood leukocytes,
 predominantly in eosinophils and granulocytes, and at lower levels
 in monocytes.
 (9)
 SIMILARITY: Contains 1 DAPIN domain.
 (10)
 SIMILARITY: Contains 1 NACHT domain.
 (11)
 SIMILARITY: Contains 8 leucine-rich (LRR) repeats.
 (12)
 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).
 (13)
 EMBL; AY095146; AAM18227.1; -
 EMBL; AY116204; AAM75142.1; -
 EMBL; AY116205; AAM75143.1; -
 EMBL; AY116206; AAM75144.1; -
 EMBL; BC028069; AAM75144.1; -
 InterPro; IPR001611; LRR.
 InterPro; IPR007091; LRR_RNinh.

phenotype including recurrent fever, cold sensitivity, sensorineural deafness, and AA amyloidosis.";

RA Arthritis Rheum. 46:2445-2452(2002).

RN [4]

RP SEQUENCE OF 391-1034 FROM N.A. (ISOFORM 1).

RC TISSUE=Blood;

RX MEDLINE=20499367; PubMed=11042152;

RA Zhang Q.-H., Ye M., Wu X.Y., Ren S.X., Zhao M., Zhao C.J., Fu G.,

RA Zhang Y., Lu G., Zhong M., Xu X.R., Han Z.-G., Zhang J.W.,

RA Tao J., Huang Q.-H., Zhou J., Hu G.X., Gu J., Chen S.-J., Chen Z.,

RT "Cloning and functional analysis of cDNAs with open reading frames for

RT 300 previously undefined genes expressed in CD34+ hematopoietic

RT stem/progenitor cells.";

RL Genome Res. 10:1546-1560(2000).

RN [5]

RP VARIANT FCAS MET-198, VARIANTS MWS ASN-303; MET-348; THR-439 AND

RX ARG-569, AND VARIANT FCAS/WMS TRP-260.

RX MEDLINE=21987640; PubMed=11992256;

RA Dode C., Le Du N., Cuisset L., Letourneur P., Berthelot J.-M.,

RA Vaudour G., Meyrier A., Watts R.A., Scott D.G.I., Nicholls A.,

RA Granel B., Frances C., Garcier F., Edery P., Boulinguez S.,

RA Domergues J.-P., Delpech M., Grateau G.;

RT "New mutations of CIAS1 that are responsible for Muckle-Wells syndrome

RT and familial cold urticaria: a novel mutation underlies both

RT syndromes.";

RL Am. J. Hum. Genet. 70:1498-1506(2002).

RN [6]

RP VARIANTS CINCA ASN-303; SER-309; ARG-358; ASN-436; SER-573 AND

RX THR-662, AND TISSUE SPECIFICITY.

RX MEDLINE=22062556; PubMed=12032915;

RA Feldmann J., Prieur A.-M., Quartier P., Berquin P., Certain S.,

RA Cortis E., Teillac-Hamel D., Fischer A., de Saint Basile G.;

RT "Chronic infantile neurological cutaneous and articular syndrome is

RT caused by mutations in CIAS1, a gene highly expressed in

RT polymorphonuclear cells and chondrocytes.";

RL Am. J. Hum. Genet. 71:198-203(2002).

CC -!- FUNCTION: May function as a potential inducer of apoptosis.

CC Interacts selectively with apoptosis-associated specklike protein

CC containing a CARD domain (ASC). This complex may function as an

CC upstream activator of NF-kappaB signaling.

CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=3;

CC Name=2;

CC IsoId=Q96P20-1; Sequence=Displayed;

CC Name=1;

CC IsoId=Q96P20-2; Sequence=VSP_005520, VSP_005521;

CC Name=3;

CC IsoId=Q96P20-3; Sequence=VSP_005519;

CC -!- TISSUE SPECIFICITY: Expressed in blood leukocytes. Strongly

CC expressed in polymorphonuclear cells, undetectable or expressed

CC at a lower magnitude in B and T lymphoblasts, respectively. High

CC level of expression detected in chondrocytes. Low or no expression

CC in the other tissues tested.

CC -!- DISEASE: Defects in CIAS1 are a cause of familial cold

CC autoinflammatory syndrome (FCAS), commonly known as familial cold

CC urticaria. FCAS is rare autosomal dominant systemic inflammatory

CC disease characterized by episodes of rash, arthralgia, fever and

CC conjunctivitis after generalized exposure to cold.

CC -!- DISEASE: Defects in CIAS1 are a cause of Muckle-Wells syndrome

CC (MWS), a rare autosomal dominant fever syndrome with episodic

CC urticaria, arthralgia, amyloidosis and progressive sensorineural

CC deafness.

CC -!- DISEASE: Defects in CIAS1 are the cause of chronic infantile

CC neurologic cutaneous and articular syndrome (CINCA), also known as

CC 'neonatal onset multisystem inflammatory disease,' or NOMID, a

CC rare congenital inflammatory disorder characterized by a triad of

CC neonatal onset of cutaneous symptoms, chronic meningitis, and

CC joint manifestations with recurrent fever and inflammation.

CC -!- SIMILARITY: Contains 1 DAPIN domain.

CC -!- SIMILARITY: Contains 1 NACHT domain.

CC -!- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.

CC -!- CAUTION: Ref.4 sequence differs from that shown due to frameshifts

CC in positions 893, 918 and 926.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AF410477; AAL33908.1; -

DR EMBL; AF427617; AAL33911.1; -

DR EMBL; AY051117; AAL12497.1; JOINED.

DR EMBL; AY051112; AAL12497.1; JOINED.

DR EMBL; AY051113; AAL12497.1; JOINED.

DR EMBL; AY051114; AAL12497.1; JOINED.

DR EMBL; AY051115; AAL12497.1; JOINED.

DR EMBL; AY051116; AAL12497.1; JOINED.

DR EMBL; AY056059; AAL12497.1; JOINED.

DR EMBL; AY056060; AAL12497.1; JOINED.

DR EMBL; AY051117; AAL12498.1; -

DR EMBL; AY051112; AAL12498.1; JOINED.

DR EMBL; AY051113; AAL12498.1; JOINED.

DR EMBL; AY051114; AAL12498.1; JOINED.

DR EMBL; AY051115; AAL12498.1; JOINED.

DR EMBL; AY051116; AAL12498.1; JOINED.

DR EMBL; AF420469; AAL65136.1; -

DR EMBL; AF468532; AAL78632.1; ALT INIT.

DR EMBL; AV092033; AAM14669.1; ALT_INIT.

DR EMBL; AF418985; AAL14640.2; ALT_INIT.

DR EMBL; AF054176; AAC39910.1; ALT_FRAME.

DR Genew; HGNC:16400; CIAS1.

DR MIM; 606416; -

DR MIM; 120100; -

DR MIM; 191900; -

DR MIM; 607115; -

DR GO; GO:0016506; P:apoptosis activator activity; NAS.

DR GO; GO:0006917; P:induction of apoptosis; NAS.

DR GO; GO:0006954; P:inflammatory response; IMP.

DR GO; GO:0007165; P:signal transduction; NAS.

DR InterPro; IPR001611; LRR.

DR InterPro; IPR007091; LRR_RNinh.

DR InterPro; IPR003590; LRR_RNinh sub.

DR InterPro; IPR007111; NACHT_NTFase.

DR InterPro; IPR004020; PAAD_DAPIN_dom.

DR Pfam; PF00560; LRR; 2.

DR SMART; SM00368; LRR_RI; 3.

DR PROSITE; PS50824; DAPIN; 1.

DR PROSITE; PS50837; NACHT; 1.

DR Apoptosis; Repeat; Leucine-rich repeat; Alternative splicing;

KW Disease mutation; Deafness.

FT DOMAIN 1 91

FT DOMAIN 218 534

FT REPEAT 738 761

FT REPEAT 795 818

FT REPEAT 852 875

FT REPEAT 881 904

FT REPEAT 909 937

FT REPEAT 939 961

FT REPEAT 966 989

FT DOMAIN 688 695

FT VARSPLIC 718 1034

FT MISSING (in isoform 3).

FT FTID=VSP_005519.

FT MISSING (in isoform 1).

FT FTID=VSP_005520.

FT MISSING (in isoform 1).

FT FTID=VSP_005521.

FT V -> M (IN FCAS AND MWS).

FT FTID=VAR_013227.

FT R -> W (IN FCAS AND MWS).

FT FTID=VAR_014104.

FT D -> N (IN CINCA AND MWS).

FT FTID=VAR_014105.

FT L -> P (IN FCAS AND MWS).

[illegible]

QY 868 TELDLISFNVLTDAGAKHLICORLPQSKLORLQVSCGLTSDCCODLASVLSPSLKEL 927
 DB 1011 THLSLSMNVFENGKLLCEVWREPSCHQDLVWKCHTAAACCSLSVCVRSRHLKSL 1070
 QY 928 DLQONLNDVGVRLCEGLRHCPACKLIRLGLPQTTLSDEMROELRALEOKPQLLIFSRR 987
 DB 1071 DLTDNALGDDGVAALCEGLKQKNSVLRLGLKACGLTSCCEALS- 1123
 QY 988 KFSVWPTPTEGLDTGENMSNTSLKQRLGSEARAASHVAOANUKLLDVSK-IPF-IAETA 1044
 DB 1124 LLSL-----NLVQNNFSPKGMKLCGSAFAC---PTSNLQIIGLWKQVFPVQIRKLL 1171
 QY 1045 ESSPPEVVPVELLCVSPASQGDHLTKPLGTDDDFW 1080
 DB 1172 EE-----VQLL-KPRVVIDGSWHS--FDEDRYW 1197

RESULT 7
 NAL2 HUMAN
 ID NAL2 HUMAN STANDARD; PRT; 1062 AA.
 AC Q9NX02; Q9BVN5; Q9H6G6; Q9HAV9; Q9NWK3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE NACT- LRR- and PYD-containing protein 2 (Nucleotide-binding site
 DE protein 1) (PYRIN-containing APAF1-like protein 2).
 GN NALP2 OR NBS1 OR PYPAF2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RN SEQUENCE FROM N.A. (ISOFORM 2).
 RP MEDLINE=21169419; PubMed=11270363;
 RA Bertin J., Distefano P.S.;
 RT "The PYRIN domain: a novel motif found in apoptosis and inflammation
 RT proteins";
 RL Cell Death Differ. 7:1273-1274 (2000).
 RN [2]
 RN SEQUENCE FROM N.A. (ISOFORM 1).
 RP MEDLINE=21148093; PubMed=11250163;
 RA Martinon F., Hofmann K., Tschopp J.;
 RT "The pyrin domain: a possible member of the death domain-fold family
 RT implicated in apoptosis and inflammation";
 RL Curr. Biol. 11:R118-R120 (2001).
 RN [3]
 RN SEQUENCE FROM N.A. (ISOFORM 1).
 RP MEDLINE=22162427; PubMed=12019269;
 RA Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,
 RA Lora J.M., Geddes B.J., Briskin M., Distefano P.S., Bertin J.;
 RT "PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates
 RT activation of NF-kappa B and caspase-1-dependent cytokine
 RT processing";
 RL J. Biol. Chem. 277:29874-29880 (2002).
 RN [4]
 RN SEQUENCE FROM N.A. (ISOFORM 1).
 RP TISSUE=Colon, and kidney epithelium;
 RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
 RA Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T.,
 RA Nakamura Y., Isoigai T., Siganos S.;
 RT "NEDO human cDNA sequencing project";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RN SEQUENCE FROM N.A. (ISOFORM 1).
 RP TISSUE=Lung, and placenta;
 RC MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klauener R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., McQuellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Boek S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Griewood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -!- FUNCTION: May be implicated in apoptosis (By similarity).
 CC -!- COFACTOR: Binds ATP (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q9NX02-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q9NX02-2; Sequence=VSP 005522;
 CC -!- SIMILARITY: Contains 1 DAPIN domain.
 CC -!- SIMILARITY: Contains 1 NACT domain.
 CC -!- SIMILARITY: Contains 9 leucine-rich (LRR) repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF298547; AAG15253.1; ALT_INIT.
 CC EMBL; AF310106; AAG30289.1; -
 CC EMBL; AF464764; AAL69962.1; -
 CC EMBL; AK000517; BAA91223.1; -
 CC EMBL; AK025952; BAB15293.1; -
 CC EMBL; AK000784; BAA91377.1; ALT_INIT.
 CC EMBL; BC003592; AAH03592.1; -
 CC EMBL; BC001039; AAH01039.1; -
 CC HSSP; P10775; 2BNH.
 CC InterPro; IPR007091; LRR_RNinh.
 CC InterPro; IPR007111; NACT_NTFase.
 CC InterPro; IPR004020; PAAD_DAPIN_dom.
 CC PROSITE; PS50824; DAPIN; 1.
 CC PROSITE; PS50837; NACT; 1.
 CC Apoptosis; ATP-binding; Leucine-rich repeat; Repeat;
 CC Alternative splicing.
 CC DOMAIN 1 94 DAPIN.
 CC DOMAIN 207 526 NACT.
 CC REPEAT 467 491 LRR 1.
 CC REPEAT 622 645 LRR 2.
 CC REPEAT 754 777 LRR 3.
 CC REPEAT 810 832 LRR 4.
 CC REPEAT 839 862 LRR 5.
 CC REPEAT 867 890 LRR 6.
 CC REPEAT 924 947 LRR 7.
 CC REPEAT 981 1005 LRR 8.
 CC REPEAT 1010 1033 LRR 9.
 CC NP_BIND 213 220 ATP (POTENTIAL).
 CC DOMAIN 518 523 POLY-GLU.
 CC VARSPLIC 133 154 Missing (in isoform 2).
 CC CONFLICT 1 1 M -> V (IN REF. 1).
 CC CONFLICT 35 35 L -> P (IN REF. 1).
 CC CONFLICT 304 304 I -> S (IN REF. 4; BAB15293).
 CC CONFLICT 364 364 R -> K (IN REF. 1).
 CC CONFLICT 980 980 MISSING (IN REF. 1).
 CC CONFLICT 1052 1052 A -> E (IN REF. 5).
 CC SEQUENCE 1062 AA; 120514 MW; 4DBB0F6E9C2BC9A7 CRC64;

Query Match		12.6%;	Score 948;	DB 1;	Length 1062;
Best Local Similarity		27.8%;	Pred. No. 6.4e-53;		
Matches		302;	Conservative 167;	Mismatches 389;	Indels 228; Gaps 32;
QY	148	WREISASLLYQALPSSPDHESQSPNAPTAVLGSWGSPPQSLAPREDEAGTQWP	207		
DB	67	WVEMASLQVFEK-----HRMDSERAKDEVREAAKLSFNKRPKPLSLGTRKERP-----P	117		
QY	208	LDTSIYITEIREREREKSEKGRPPAAVVGTPPOAHTSLQPH-----HHPEWPSVRES	262		
DB	118	LD-----VDEMLEREKTEQAFTETKGNVICLGEVFKGKPKDKNRCRYILTKFERE-	170		
QY	263	LCSTWPKWKNEDFN---OKFTQTLQLLQRPHPRSQDPLVGRSWPDYVEENRGLHIEIRDLFG	319		
DB	171	MNKSPPGDSKEQVVAERYKMLIPFNP-----RVLPQ	203		
QY	320	PGLDTQEPRIVILQAGAGIKSTLAROQVEANGRGOLYGDRTQHVFPYFSCRELAOSKVV	379		
DB	204	PFSYT-----VVLGYPAGLGKTTLAQKMLDWAEDNLI-HKFKYAFYLSRSLSLGPGS	257		
QY	380	LAELGKDGCTATPAPTRQILSRPERLLFILDGVDE-----PGWVLOEPPSELCLHWSQOP	435		
DB	258	FIELFRDWELODDIPHLAQARKILFVIDGDELGAAPGALIE-----DICDWEKKP	313		
QY	436	ADALLGSLGKTLTPASFLITARTALQNLIPSLQEARVVEVLGFSSESRKEYFYRYFT	495		
DB	314	VPVLLGSLNRYMLPKAALLVTRPRALRDLRLAIEEPIYIRVEGFLEEDRRAYFLRHFG	373		
QY	496	DERQAIRFRLVKSNEKELWALCLVPVWSWLAETCLMQMKREKELTSTKTTTTLCLHVL	555		
DB	374	DEDQAMRAFELMRSNAALFQLSAPAVCVIVCTTTLQMEKGEDPVPTCLTRTGLFLREL	433		
QY	556	AQAL-QAQLGQLRDLCSLAEGIQKTLSPDRLRKHGLDGAISTFLKMGILQHP	614		
DB	434	CSRFPOAQLRGALRTLSLLAQGLWAQTSVLHREDLERLGQVSDRLFLDGLDILQR	493		
QY	615	IPLS-YSFHLCQFEEFAAMSYYL--EDEKGRGKSNCTIDLEKTEAYGIGHLFGASIT	671		
DB	494	VSKGCVSFHLSFQFLTALFYLEKEBEDRDGHTWDIGVQKLS--GVERLRNPDL	551		
QY	672	R---FLLGLSDEGEEMENIFHRLSQGRNLMQWVPSL-QLLQ-----PHSL----	716		
DB	552	QAGYVSFGLANERAKRELEATFCRMS-----PDIKQELLRCDISCKGHSVTDL	602		
QY	717	ESLHCLYETRNKFTLTQWMAHEMGMCVETDMELLVCTCFKFSRHVKLQL-----	769		
DB	603	QELLGLYSQEBELKVEVNAQKETSLLHNA-VVVVPSFVCVKHCRNLRQKMSLQVIKEN	661		
QY	770	IEGRQ-----IEGRQ-----HRSTWSPTMV-----	784		
DB	662	LPENVTASBDAEVSQDDQHMLPFWTDLCSIFGSKNDLGLAINDSFLSASLVRLCE	721		
QY	785	-----VLFRRVVPVDAWQILFSLVKVTRNLKELDLSGNSLSHSAVKSCKTLR	833		
DB	722	QIASDTCRLQVVVFNKISPADAHNLCIA-LRGHKTVTYTLTQND-QDMFPALCEVLR	779		
QY	834	RPCILLETILACCGLTAECDKDLAFLGRANQTLTDLDSFNVLTADGAKHLQRLRQPS	893		
DB	780	HPECNRYLGLVSCSATTOQWADLSLALEVNSQSLTCVNLSDNELLDEGAKLYTLRHPK	839		
QY	894	CKLQRLQVSCGLTSDDCCLASVLSASPSLKELDLQONLDDVGVRLLCGLRHCPACKL	953		
DB	840	CFIQLRSLNCHLITENCKDLAAVLVVSRELTHLCLAKPIGNTGVKFLCEGLRYPECKL	899		
QY	954	IRLGLDQTLTSDMRQELRALEQKQLLIFSRKPSVMTPTBGLDTGMSNSTSLKQ	1013		
DB	900	QTLVLANCDITSDGCCDLTKLQEKSSLLC-----LDLG-----	933		
QY	1014	RLGSEAAHVAQANIKLLDVSKIFPIAEI-----AESSEPVVPVLLCVPSASQ----	1066		
DB	934	-----LNHIGVKMGKFLCEALRKLCLNCLWLWGCSTIPPFSCEDLCSALSCNQLVT	986		

QY	1067	DLHTKPLGTD-----DDFWGPTGPVATEVVDVKEKNLY---R	1099
DB	987	LDLQGNPLGSSGVKMLFETLTCSSGTLRLTKIDDFNDELKLEEEIEKNPQLIIDTE	1046
QY	1100	VHFPVA 1105	
DB	1047	KHPWA 1052	
RESULT 8			
ID	PYA3	HUMAN	
AC	Q8WX94;	STANDARD;	PRT; 980 AA.
DT	28-FEB-2003	(Rel. 41, Created)	
DT	28-FEB-2003	(Rel. 41, Last sequence update)	
DT	28-FEB-2003	(Rel. 41, Last annotation update)	
DE	PYRIN-containing APAF1-like protein 3.		
GN	PYPAF3.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=22162427; PubMed=12019269;		
RA	Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,		
RA	Lora J.M., Geddes B.J., Briskin M., DiStefano P.S., Bertin J.;		
RT	"PYPAF7, a novel PYRIN-containing APAF1-like protein that regulates		
RT	activation of NF-kappa B and caspase-1-dependent cytokine		
RT	processing.";		
RL	J. Biol. Chem. 277:29874-29880(2002).		
CC	- SIMILARITY: Contains 1 DAPIN domain.		
CC	- SIMILARITY: Contains 1 NACHT domain.		
CC	- SIMILARITY: Contains 9 leucine-rich (LRR) repeats.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
EMBL:	AF464765; AAL69963.1; -		
DR	InterPro; IPR007091; LRR_RNinh.		
DR	InterPro; IPR007111; NACHT_NTPase.		
DR	InterPro; IPR004020; PAAD_DAPIN dom.		
DR	Pfam; PF02758; PAAD_DAPIN; 1.		
DR	PROSITE; PS50824; DAPIN; 1.		
DR	PROSITE; PS50837; NACHT; 1.		
KW	ATP-binding; Leucine-rich repeat; Repeat.		
FT	DOMAIN	1	93
FT	DOMAIN	172	491
FT	REPEAT	614	638
FT	REPEAT	674	697
FT	REPEAT	760	784
FT	REPEAT	788	810
FT	REPEAT	817	840
FT	REPEAT	845	868
FT	REPEAT	874	897
FT	REPEAT	902	928
FT	REPEAT	933	957
FT	NP_BIND	178	185
SQ	SEQUENCE 980 AA; 111806 MW; 822AF2FD4338003D CRC64;		
Query Match 12.1%; Score 908.5; DB 1; Length 980;			
Best Local Similarity 30.2%; Pred. No. 2e-50;			
Matches 261; Conservative 128; Mismatches 284; Indels 191; Gaps 21;			
QY	294	PLVKRSWPDYVEEN-----RGLHIEIRD-----LFGPGLDTQSPR	332
DB	117	PGEKSGWRNSMEKQSLVWKNFTWQGDIDFNHDDVTLRNQRFIFPLNPRTPKLTPTVVL	176


```

QY 241 PQAHTSLQPHHPWSPVRESLCSTWPKNEDFNQKFTQLLLQRPSPQDPLVGRSW 300
Db 179 ---THI-----FNRLPR-----188
QY 301 PDVVENRHLEIRDLFGGLDQEPRIIVILQGAAGICKSTLQVKEAWRGQYGLYDGR 360
Db 189 ---DEGR-----RPLTVLQGPAGIGIKTAAKILYDMAGKLYQOQ 228
QY 361 FQHVFFSCRELAQSK-VVSLAELIGKDGATAPAPIRQILSPERLLFLDQWDE-PGWV 418
Db 229 VDFAFMPGCELLERPTSLADLILDQCPDGAQVPMQAPQRLFLDGADELPA--A 286
QY 419 LOEPPSELCHWSQOPADALGLSLGKTLTLPASFLITARTALQNLPLPSLEQARWVEV 478
Db 287 LGSPEAAPCTDPEASAGARVLGGLSKALLPTALLVTTTAAAPQRLQGLRCSQCAEV 346
QY 479 LGFESSREYFVRYPTDQRAIRAPLVKSKELWALCLVPMVSWLACTCLMQMKRKE 538
Db 347 RGFSDKDKKXYFKFPRDERRAFRYFVKNETLFCALCFVPFCVIVCTVLRQQLRGR 406
QY 539 KLTLSKTTTTLCHVLAQALQAPL--GPO-----LRDLCSLAAGIMOKKTLFGPDLLR 592
Db 407 DLSRTSKTTTIVLLFITSVLSAPVADGPRLQGLDLNLCRLAREGLVGRRAQFAKELE 466
QY 593 KHGLDGAIST-FLKM-----GILQEHPIPLSYFHLFCQEPFAAMSYVLED-----E 640
Db 467 QLELRGSKVQTLFLSKKELPGVLETE--VTYQIDQSFQFLAALSYLLEDGGVPRTA 523
QY 641 KGRG-----KHSCNIIDLEKTLBAYGHLFGASTTRFLGLLSDRGEREMENIFHC 692
Db 524 GGVGTLLRGDAOPSHLVL-----TTRFLGLLSAERMDIERHFQCG 565
QY 693 RLSQ--GRNLMQWV-----PSLQLLQP-----HSLSLHCL 722
Db 566 MVSERVKQALRWVQOQGCQCPGAVEVTEGAKLEDTEPEEBEGBEPNPLELLYCL 625
QY 723 YETRNKTFITQVMAHFEEMGM-----CVETDMELLVCTFCIKFSRHVKKLQIEGRQRS 777
Db 626 YETQEDAFVQALCRFPALQVRFC--RMDVAVLSYVCRCCPAGQALRLISCR---- 678
QY 778 TWSPTMVVLFWRVVPVTDAYWQILFSLVKVTRNL---KELDLSGNSLSHSAVSKLCKTLR 834
Db 679 -----LVAQEKKKKSLGKGLQASLGSGSSQTTKQLPASLH 717
QY 835 -----PRCLETLRACGLTAEDCKDLAFGLRANQTLFELDSLFNVLTDAGAKHLC 886
Db 718 PLFQAMTDPLCHLSSUTLSHCKLPDVAVCRLDSLALRAAPALTELGLLHRLSEAGLRML 777
QY 887 QRLRQPSCKLQ--RLQL--VSCGLTSDCCQDLASVLSASPSLKELDLQNNLDDVGVRL 942
Db 778 EGLAWPQCRVQTVRVQLPDQRLG-----QYLVGMLRQSPALTTLDLSCQQLPAPMVTYL 832
QY 943 CEGRLHPACKLRLGLDQDTLSDQMQLRALEQERQPLLI 983
Db 833 CAVLQHQGGGLQTLSLASVELSSEQLQELQAVKRAKPDVLI 873

RESULT 10
PYAS_RAT
ID PYAS_RAT STANDARD; PRT; 854 AA.
AC Q63035;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE PYRIN-containing APAFI-like protein 5-like (Angiotensin II/vasopressin
DE receptor).
GN PYAP5 OR AVR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]

```

```

RP CONCEPTUAL TRANSLATION OF 1-343.
RA Hinz U.;
RL Unpublished observations (FEB-2003).
RN [2]
RP SEQUENCE OF 344-854 FROM N.A., MUTAGENESIS OF GLU-393, AND FUNCTION.
RC TISSUE=Kidney;
RX MEDLINE=96071640; PubMed=7489366;
RA Ruiz-Opazo N., Akimoto K., Herrera V.L.M.;
RT "Identification of a novel dual angiotensin II/vasopressin receptor on
RT the basis of molecular recognition theory.";
RL Nat. Med. 1:1074-1081(1995).
RN [3]
RP VARIANTS SALT-SENSITIVE HYPERTENSION SER-492 AND ARG-536.
RX MEDLINE=21980185; PubMed=11984003;
RA Ruiz-Opazo N., Lopez L.V., Herrera V.L.M.;
RT "The dual AngII/AVP receptor gene N119S/C163R variant exhibits
RT sodium-induced dysfunction and cosegregates with salt-sensitive
RT hypertension in the Dahl salt-sensitive hypertensive rat model.";
RL Mol. Med. 8:24-32(2002).
RN [4]
RP IDENTIFICATION OF MAMMALIAN ORTHOLOGS OF PYAP5.
RA Albrecht M., Domingues P.S., Schreiber S., Lengauer T.;
RT "Identification of mammalian orthologs associates PYAP5 with distinct
RT functional roles.";
RL FEBS Lett. 538:173-177(2003).
CC -!- FUNCTION: May mediate activation of CASP1 via ASC and promote
CC activation of NF-kappa-B (By similarity). Angiotensin II and
CC vasopressin binding protein. May stimulate cAMP accumulation.
CC -!- SUBUNIT: Binds to ASC with its DAPIN domain (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- TISSUE SPECIFICITY: Highly expressed in kidney. Detected at low
CC levels in all tissues tested.
CC -!- DISEASE: Defects in PYAP5 may be a cause of salt-sensitive
CC hypertension.
CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- SIMILARITY: Contains 4 leucine-rich (LRR) repeats.
CC -!- CAUTION: The N-terminus was extended using ESTs and genomic
CC sequences, in analogy to ortholog sequences.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M85183; AAA03623.1; ALT_INIT.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh.
DR Pfam; PF00560; LRR; 2.
DR Pfam; PF02758; PAAD_DAPIN; 1.
DR PROSITE; PS50824; DAPIN; 1.
DR PROSITE; PS50837; NACHT; 1.
KW ATP-binding; Leucine-rich repeat; Repeat.
FT DOMAIN 1 103 DAPIN.
FT DOMAIN 168 484 NACHT.
FT DOMAIN 564 604 ASP/GLU-RICH.
FT DOMAIN 655 662 POLY-LYS.
FT REPEAT 433 458 LRR 1.
FT REPEAT 610 633 LRR 2.
FT REPEAT 722 745 LRR 3.
FT REPEAT 812 836 LRR 4.
FT NP_BIND 174 181 ATP (POTENTIAL).
FT VARIANT 492 492 N -> S (IN SALT-SENSITIVE HYPERTENSION).
FT VARIANT 536 536 C -> R (IN SALT-SENSITIVE HYPERTENSION).
FT MUTAGEN 766 766 E->K; ABOLISHES ANGIOTENSIN II BINDING.
SQ SEQUENCE 854 AA; 95292 MW; D7BBD922D77B3734 CRC64;

Query Match 11.9%; Score 895; DB 1; Length 854;
Best Local Similarity 33.2%; Pred. NO. 1.2e-49;
Matches 261; Conservative 118; Mismatches 298; Indels 110; Gaps 22;

```


FT REPEAT 640 664 LRR 1.
 FT REPEAT 691 714 LRR 2.
 FT REPEAT 737 762 LRR 3.
 FT REPEAT 776 799 LRR 4.
 FT REPEAT 804 832 LRR 5.
 FT REPEAT 834 856 LRR 6.
 FT REPEAT 861 884 LRR 7.
 FT REPEAT 890 913 LRR 8.
 FT REPEAT 918 941 LRR 9.
 FT REPEAT 950 974 LRR 10.
 FT REPEAT 975 1003 LRR 11.
 FT REPEAT 1005 1027 LRR 12.
 FT REPEAT 1032 1059 LRR 13.
 FT NP_BIND 197 204
 FT VARIANT 7
 FT VARIANT 42 42 E -> D (IN STRAIN 129/SV).
 FT VARIANT 87 87 G -> L (IN STRAIN 129/SV).
 FT VARIANT 256 256 R -> K (IN STRAIN 129/SV).
 FT VARIANT 977 977 L -> P (IN STRAIN 129/SV).
 FT VARIANT 1003 1004 H -> Q (IN STRAIN 129/SV).
 FT SEQUENCE 1111 AA; 125502 MW; 4016A5D67A1C01F4 CRC64;
 Query Match 11.7%; Score 882; DB 1; Length 1111;
 Best Local Similarity 26.1%; Pred. No. 1.2e-48;
 Matches 274; Conservative 154; Mismatches 350; Indels 272; Gaps 25;

QY 162 SSPDHPSPQESNPAP-----TSTAVLGSWGSPPPQS-----LAPR-----BOEAPGTQWPLD 209
 DB 28 TSPENDSKSIQKQDQEQEQTSESTWG-----PPEKDSKAILKARGLEEBEQKSESTWSPSE 83
 QY 210 ETSGIYYTIRERERKSEKRPAAVGTPTQANTSLQPHHPWEPSPVRESLSTWPM 269
 DB 84 NVSRAILKSGSEVEOASERK-----MTSPENDSKSIQKQDQEQEQTSETLQS----- 133
 QY 270 KNEDFNQKQTQLLLQRPSPQDPLVKRSPWPVVEENRGHLE----- 313
 DB 134 KEED-----EYTE-----ADKONGGD-LQDYKAHVIAKPTSDVLHYDSPENK 175
 QY 314 -IRDLFGPLDTEPRIVLQGAAGIKSTLARQVKEAWGRGOLYGDGFQHFVYFSCREL 372
 DB 176 LISDAFKPQKTQPTHTILHGRPGVGSALARSIVLGHQAQKLP-QKMSFVIFFSVRRI 234
 QY 373 AQSKVSVLAELIGKGTATPAPIROILSRPERLLFDLGDVDEPGWLOPSPSELCHLWQ 432
 DB 235 KWTEKSLAQLIAKECPDSWDLTKMSQPERLLFVIDGLDDWDSVLQDHDWTLSDRWDK 294
 QY 433 PQADALLGSLKLTLPASFLITARTALQNLIFSLQEARWVVLGFSSESRSRKEYFYR 492
 DB 295 EQPIYILMYSLRLKALLPOSFLITRTNTGLEKLKSNVWVSPLYILVEGLSASRRSOLVLE 354
 QY 493 YFTDERQATRAPRLVKSNEKLMALCLVPMVSWLACTLMOQMKRKEKLTITSKTTT----- 548
 DB 355 NISNESDRIOVFHSLNENHOLFQOQAPSVCVSLVCEALQKLGKRCRTLPCTLTGLYA 414
 QY 549 TLCLHVL-----AQAQAQPLGQPLRDLCSLAAEGIWQKTLFSPDRLKRGHLDGAIIST 603
 DB 415 TLVHFQTLTKRPSQSALSQBEQITVGLCMAAEGVWTVRESVYDLDLKNYSKESEILA 474
 QY 604 FLKMGILQE--HIPISYSYTIHCFQEFPAAMSVYLEDEKGRKH-----SNCIIDLEK 655
 DB 475 LFHMNILLQVGHNSQCYVFSHLSLQDFFAALYVYLEGEEMWQHFCFIENQKRSIMEVR 534
 QY 656 TLEAYGHLFGASTTRELGLLSDSEGEREMENI FHCRL--SQRNLMOWPSLQLLQ 713
 DB 535 TDDT-----RULGMK--RFLFGLMNKQILKTLVLFYFVPTVEQKQHWV-----SLIAQQ 585
 QY 714 -----HSLSEHCLYETRNKTLFQVMAHPEMGMCVETDMELLVCTFCIKFSRHVKK 766
 DB 586 VNGTSPMDTLDAICYLFESQDEFFVGALKRFQEVWLLINQKWDLVKSVYCLKHQCNLKA 645
 QY 767 LQL-----IEGRQHRSTWSPMVLV-FRWVPVTDAYWQILFVSLVKVTRNLKELDLGNSLS 821
 DB 646 IRVDIRDLISVDNTLCLPQVTVQETQCKPLLMWGNFCSVLGSLRNKELDLGDSILS 705

QY 822 HSAVKS----- 828
 DB 706 QRAMKILCLELRNOSCRIOKLTFFKSAEVVSLGLKHLWKLKLLFSNQNLKYLNLGNTPMKDDDM 765
 QY 829 ---CKTLRPRCLLETLR----- 843
 DB 766 KLACEALKHPKCSVETLRLDSCELTIIQYEMISTLLITRRLKCLSLAKNRVGVKSMISL 825
 QY 844 -----LAGGLTAEDCKDLAFGLRANQTLTDLTSLFNVLTDAGAKHLQOR 888
 DB 826 GNALSSWCLLQKLLDNCGLTPASCHLLVLSALFSNQNLTHLCLSNNSLGTGVQOLCOF 885
 QY 889 LRQPSCKLQ----- 898
 DB 886 LRNPECALQRLTLNHCNIVDDAYGFLAMLANNTKLTHTLSLTMTNVPVGDGAMKLLCEALKE 945
 QY 899 -----LQVSCGLTSCCCQDLASVLSASPSLKELDLQONNLDVGVRLLCGLRHPAC 951
 DB 946 PTCYLQLELVDCQLTQCCEDLACMTTTRTKLKSLLDGNALGDKGVITLCEGLKQSS 1005
 QY 952 KLIRGLQDTLLSDMRQELRALEQEKPOL 981
 DB 1006 SURRLGLGACKLTSNCCREALSLAISCNPHL 1035

RESULT 12
 PYAS MOUSE
 ID PYAS MOUSE STANDARD; PRT; 843 AA.
 AC Q91WS2; Q8KOL4;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE PYRIN-containing APAF1-like protein 5-like.
 GN PYPAF5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1] _TaxID=10090;
 RP CONCEPTUAL TRANSLATION OF 1-287.
 RA Hinz U.;
 RL Unpublished observations (FEB-2003).
 RN [2]
 RC SEQUENCE OF 288-843 FROM N.A.
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Hellon E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smailus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [3]
 RP IDENTIFICATION OF MAMMALIAN ORTHOLOGS OF PYPAF5
 RA Albrecht M., Domingues F.S., Schreiber S., Lengauer T.;
 RT "Identification of mammalian orthologs associates PYPAF5 with distinct
 RT functional roles.";
 RL FEBS Lett. 538:173-177 (2003).


```

CC TISSUE=Glial tumor;
RA Kawakami B., Sugiyama A., Takemoto M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,
RA Isomoto Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Muraoka K., Kanehori K., Takahashi-Fujii A., Oshima A.,
RA Suzuki Y., Sugano S., Nagahara K., Masuho Y., Nagai K., Isogai T.;
RL "NEDO human cDNA sequencing project.";
CC Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P59045-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P59045-2; Sequence=VSP_007068;
CC Note=No experimental confirmation available;
CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AY095145; AAM14632.1; -
CC EMBL; BC034730; AAH34730.1; -
CC EMBL; AK090621; BAC03490.1; ALT INT.
CC InterPro: IPR007091; LRR_RNinh.
CC InterPro: IPR003590; LRR_RNinh sub.
CC InterPro: IPR007111; NACHT_NTPase.
CC InterPro: IPR004020; PAAD_DAPIN_dom.
CC Pfam; PF00560; LRR; 1.
CC Pfam; PF02758; PAAD_DAPIN; 1.
CC SMART; SM00368; LRR_RI; 9.
CC PROSITE; PS50824; DAPIN; 1.
CC PROSITE; PS50837; NACHT; 1.
CC ATP-binding; Nucleotide-rich repeat; Repeat; Alternative splicing.
KW DOMAIN 1 91 DAPIN.
FT DOMAIN 147 470 NACHT.
FT REPEAT 588 611 LRR 1.
FT REPEAT 632 655 LRR 2.
FT REPEAT 745 768 LRR 3.
FT REPEAT 802 827 LRR 4.
FT REPEAT 859 882 LRR 5.
FT REPEAT 919 944 LRR 6.
FT NP_BIND 153 160 ATP (POTENTIAL).
FT VARSPIC 614 668 SOMKSLVYWRICSLFVTWESLRELHIFDNLNGISERILS
KALEHSCKLRLTK -> R (in isoform 2).
FT FTId=VSP_007068.
FT T -> I (IN REF. 3).
FT CONFLICT 788 788
FT SEQUENCE 1033 AA; 117794 MW; A28880485FBA49AC CRC64;

Query Match 9.4%; Score 711; DB 1; Length 1033;
Best Local Similarity 26.7%; Pred. No. 1.le-37;
Matches 229; Conservative 126; Mismatches 294; Indels 208; Gaps 20;

QY 330 VILQAGAGICKSTLQKVEAGRGQLYGDFQHFYFYSFCELAQSKVSLAEILKDGST 389
Db 149 VFLMGERASGKTIVINLAVLRWIKGEWQNMISVYVHLTSHEINQMTNSSLAEILAKDWP 208
QY 390 ATPAPTRQILSRPERLLFLDGVDEPGCVLQEPSSBELCLHWSQPADALLGSLGKTL 449
Db 209 DGQAPIADILSDPKLLFILELDNIRFELNVNSESALCSNSTQKVPVLLVSLKRWMA 268
QY 450 PEASFLITARTALQNLIPLEQARVVEVLGFSESSRKEFYFYFDEQATRAFLVKS 509
Db 269 PCGWFLLISRPRTGNVKTFLKEVDCCTTLQJNSNGKREIFYNFFPKDRQRASALQLVHS 328
QY 510 NKEWALCLVPWVSWLACTCLMQMKRKBKLTLSKTTTTLCLHLAQLAQAP----- 563

```

```

Db 329 DEILVGLCRVAILCTVILKQMDKGRDFQLCQQTPTDLHAFLADALTSEAGLTANQ 388
QY 564 --LGPQLDCLSLAAGIWKQKTLFSPDLRKHGDLGDAIISTFLKXGILQEHPIPLS--- 618
Db 389 YHLG--LLKRLCLLAAGLFLSTLNFSGEDLCRCVGTEDVSVLQAANIL-----LPSNTHK 443
QY 619 --YSFIHLCFQFFFAAMSIVL-----EDEKGRGKHSNCIIDELEKTLEAYGIHG 664
Db 444 DRYKTHLNVQEFCTAIAFLMAVNPYLIPSGSREYKEKREQYSD----- 487
QY 665 LFCASTTRFLGLLSDEGEREMENTPHCELSGOGRNLMQWVPSLQLLLQ-----PHSLES 718
Db 488 --FNQVET--FIFGLNANRRKILETSFGYQLPMDVDFKWSYGVYMKHLDRDPKLTTHMPL 545
QY 719 LHCLYETRNKTLTQVMAHFEEMGCMVETDELLVCTFCIKFSRHVKKLQLEGR--QHR 776
Db 546 FYCLYENREEFVKTIIDALMEVTYVLQSDKDMVSVLYCLDYCCHRLTLKLSVQRFQNK 605
QY 777 STW-SPTMVLFPRWVPVTDAYWQILFSLVKVTRNLKEL-----DLSG-----NSLSHS 823
Db 606 EPLIRPTASQMSLV-----YWREICSLFYTWESLRELHIFDNLNGISERILSKALEHS 660
QY 824 AVK-----SLCKTLRR-----PRCLL 839
Db 661 SKLRLTKLSYVSTASGFEDLLKALARNRSLTYLSINCTSLNMFSLHLHLHEPTCQI 720
QY 840 ETLRLAGGLTAEDCKDLAFGLRANQTLTDLSPNVLTDAGAKHLQORLQPSCKLQRL 899
Db 721 SHLSMKCDLRASECEETASLISGSRKULTLSSNPLRSDGNWILCDALLHPNCTLISL 780
QY 900 QLVSCGLTSDCCQDLASVLSASPSLKELD----- 929
Db 781 VLVFCLLTENCSSALGRVLFPSTLRQLDLCVNRKLVYVHVTPLFPPTCQLEELHLS 840
QY 930 -----QQNLDDVGVRLLCGLRHPACKLIRLGLDQTT 962
Db 841 GCFFSSDICQYIAIVIAITNEKLSLEIGSNKTEDAGMQLCGLRHPNCLVNIQLECM 900
QY 963 LSEDMRQELRALEQEKQKPLLIFSRKPSVMTPTGLDTGEMSNSTSLKROKRGSGERAAS 1022
Db 901 LFSACCRSL-----ASVLITNKTL-----RUNLQNHLGNDGVA- 935
QY 1023 HVAQANLKLLDVSKIFF 1039
Db 936 -----KLE-SLISP 944

RESULT 14
CAR8 HUMAN STANDARD; PRT; 431 AA.
AC Q9Y2G2; Q96P82;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Caspase recruit domain protein 8 (Apoptotic protein NDPPI) (DACA)
DE (CARD-inhibitor of NF-kappaB activating ligand) (CARDINAL) (TUCAN).
GN CARD8 OR NDPPI OR KIAA0955.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RC TISSUE=Brain;
RX MEDLINE=99246063; PubMed=10231032;
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIII.
RT The complete sequences of 100 new cDNA clones from brain which code
RL for large proteins in vitro.";
RN DNA Res. 6:63-70(1999).

```


SEQUENCE FROM N.A. (ISOFORM LONG).
RA MEDLINE=21950691; PubMed=11921383;
RX Razmaria M., Srinivasula S.M., Wang L., Poyet J.-L., Geddes B.J.,
RA DiStefano P.S., Birtin J., Alnemri E.S.,
RT "CARD-8 protein, a new CARD family member that regulates caspase-1
activation and apoptosis.";
RL J. Biol. Chem. 277:13952-13958(2002).
RN [3]
SEQUENCE FROM N.A. (ISOFORM LONG).
RA Zhang H.,
RT "A novel apoptotic protein, NDP1, containing CARD and BH3 domains.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
SEQUENCE FROM N.A. (ISOFORM LONG).
RA Quiet C., Vito P.,
RT "DACAR, a novel CARD-containing protein.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
SEQUENCE FROM N.A. (ISOFORM LONG).
RX MEDLINE=21570185; PubMed=11551959;
RA Bouchier-Hayes L., Conroy H., Egan H., Adrain C., Creagh E.M.,
RA MacFarlane M., Martin S.J.,
RT "CARDINAL, a novel caspase recruitment domain protein, is an inhibitor
of multiple NF-kappa B activation pathways.";
RL J. Biol. Chem. 276:44069-44077(2001).
RN [6]
SEQUENCE FROM N.A. (ISOFORM SHORT).
RC TISSUE=Kidney;
RA Guo J.H., Yu L.,
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [7]
CHARACTERIZATION.
RX MEDLINE=21402909; PubMed=11408476;
RA Pathan N., Marusawa H., Krajewska M., Mateuszawa S.-I., Kim H.,
RA Okada K., Torii S., Kitada S., Krajewski S., Welsh K., Plo F.,
RA Godzik A., Reed J.C.,
RT "TUCAN, an antiapoptotic caspase-associated recruitment domain family
protein overexpressed in cancer.";
RL J. Biol. Chem. 276:32220-32229(2001).
RN [8]
CHARACTERIZATION, AND MUTAGENESIS OF LYS-366.
RX MEDLINE=22062958; PubMed=12067710;
RA Stilo R., Leonardi A., Formisano L., Di Jeso B., Vito P., Liguoro D.,
RT "TUCAN/CARDINAL and DRAL participate in a common pathway for
modulation of NF-kappaB activation.";
RL FEBS Lett. 521:165-169(2002).
CC -!- FUNCTION: Inhibits NF-kappaB activation. May participate in a
regulatory mechanism that coordinates cellular responses
controlled by NF-kappaB transcription factor. Involved in the
negative regulation of caspase-1.
CC -!- SUBUNIT: May form homodimers. Interacts with NEMO and DRAL. Binds
to caspase-1, pseudo-ICE and ICEBERG.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC -!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=Long;
IsoId=Q9Y2G2-1; Sequence=Displayed;
Name=Short;
IsoId=Q9Y2G2-2; Sequence=VSP_000782, VSP_000783;
CC -!- TISSUE SPECIFICITY: High expression in lung, ovary, testis and
placenta. Lower expression in heart, kidney and liver. Also
expressed in spleen, lymph node and bone marrow.
CC -!- SIMILARITY: Contains 1 CARD domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB023172; BAA76799.1; -.

DR EMBL; AF322184; AAG50014.1; -.
DR EMBL; AF331519; AAK01126.1; -.
DR EMBL; AY026322; AAK08982.1; -.
DR EMBL; AF405558; AAL02427.1; -.
DR EMBL; AF511652; AAM46959.1; -.
DR InterPro; IPR001315; CARD.
DR Pfam; PF00619; CARD; 1.
DR SMART; SM00114; CARD; 1.
DR PROSITE; PS0209; CARD; 1.
KW Apoptosis; Nuclear protein; Alternative splicing.
FT DOMAIN 340 430
FT VARSPPLIC 282 286
FT VARSPPLIC 287 431
FT MUTAGEN 366 366
FT CONFLICT 60 60
FT CONFLICT 326 326
FT CONFLICT 422 422
SQ SEQUENCE 431 AA; 48932 MW; CB54D130807732E6 CRC64;
Query Match 8.9%; Score 668; DB 1; Length 431;
Best Local Similarity 39.5%; Pred. No. 1.7e-35;
Matches 151; Conservative 53; Mismatches 122; Indels 56; Gaps 7;
QY 1079 FWGPTGPVATEVVDKEKNLYRVHPVAGSYRWPNITGLCFVNRVAVTVEIFCVMDDQLG- 1137
DB 56 FLGPEGVDVELIDKSTNRYSVWPTAGWYLSATGLGLVYRDEVTVTIATGSHSOHLAL 115
QY 1138 EINPOHSMVAGPLLDIAKEP-GAVEAVHLPFHVALQGHVDTSLSFQMAHPKBEGLLEK 1196
DB 116 DLQHEQMLVGGLPFDVTAEPPEAAVEIHLPHFISLQ-GEVDVSFWLVAHPKNEGMVLEH 174
QY 1197 PARVELHHIVLENFSPLGVLLKXIHNAIRFIPVTSVVLLYHRVHPPEVTFHLYLIPSD 1256
DB 175 PARVEFPYAVLESFSLMGILLRIASCTRLSIPITSNTLIYHHPHEDIKFHLVLPSPD 234
QY 1257 CSIR-----KELELCYRSPGSDQ 1274
DB 235 ALLTKAIDDEDRFHGVRLQTPPMEPLNFGSSIVVNSANLKVMPKELSYRSPGRIQ 294
QY 1275 LFSEPYVGHLSGIRLQVKDKDETLVWEALVKPGDLMPATTLIPPAIAPVSPDAPQL 1334
DB 295 HFSKYAGQMKPEIQLEITEKRHGTLVMDTEVKPVDLQVAAAPP-----PFSGAA- 346
QY 1335 LHFVDQYREQLIARVTSVEVLDKLH-GQVLSQSYERVLAEINTPSQMRKLFSLSQSMD 1393
DB 347 --FVKNHRQLQARMGDLKGLVLDLQDNEVLTEKEKELVEQETKRSKNEALLSVEKKG 404
QY 1394 RKCKDGLYOALKETHPHLIMEL 1415
DB 405 DLALDVLFRSISERDPYLVSYL 426
RESULT 15
CARD HUMAN
ID CARD HUMAN STANDARD; PRT; 1040 AA.
AC Q9HC29; Q96RH5; Q96RH6; Q96RH8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Caspase recruitment domain protein 15 (Nod2 protein) (Inflammatory
bowel disease protein 1).
DE bowel disease protein 1).
GN CARD15 OR NOD2 OR IBDL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), MUTAGENESIS OF LYS-305, AND
RP VARIANT ARG-908.
RC TISSUE=Breast;
RX PubMed=11087742;

RA Ogura Y., Inohara N., Benito A., Chen F.F., Yamaoka S., Nunez G.;
RT "Nod2, a Nod1/Api-1 family member that is restricted to monocytes and
RL activates NF-kappaB.";
RN J. Biol. Chem. 276:4812-4818(2001).
[2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2). VARIANTS CD THR-140; ARG-157;
RP CYS-235; ARG-248; SER-269; ASN-291; VAL-301; TRP-311;
RP VAL-348; ARG-342; CYS-373; SER-414; LEU-431; VAL-432; LYS-441;
RP VAL-612; THR-612; TRP-684; TRP-702; CYS-703; CYS-713; GLY-725;
RP VAL-755; VAL-758; MET-793; LYS-843; SER-853; VAL-863;
RP ARG-908; ASP-918; ASP-924 AND ILE-955. VARIANTS ULCERATIVE COLITIS
RP THR-140 AND THR-885, AND VARIANT MET-189.
RC TISSUE=Leukocyte;
RX MEDLINE=21279172; PubMed=11385576;
RA Hugot J.-P., Chamaillard M., Zouali H., Lesage S., Cezard J.-P.,
RA Belaiche J., Almer S., Tysk C., O'Morain C.A., Gassull M., Binder V.,
RA Finkel Y., Cortot A., Modigliani R., Laurent-Puig P.,
RA Gower-Rousseau C., Macry J., Colombel J.-F., Sahbatou M., Thomas G.;
RT "Association of NOD2 leucine-rich repeat variants with susceptibility
RL to Crohn's disease.";
RN Nature 411:599-603(2001).
[3]
RP VARIANTS BS GLN-334; TRP-334 AND PHE-469.
RX MEDLINE=21419644; PubMed=11528384;
RA Miceli-Richard C., Lesage S., Rybojad M., Prieur A.M.,
RA Manouvrier-Hanu S., Hafner R., Chamaillard M., Zouali H., Thomas G.,
RA Hugot J.-P.;
RT "CARD15 mutations in Blau syndrome.";
RL Nat. Genet. 29:19-20(2001).
CC -!- FUNCTION: Induces NF-kappaB via RICK (CARDIAC, RIP2) and IKK-
CC gamma. Confers responsiveness to intracellular bacterial
CC lipopolysaccharides (LPS).
CC -!- SUBUNIT: Binds to RICK by CARD-CARD interaction.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative initiation;
CC Comment=2 isoforms, 1/Nod2 (shown here) and 2/Nod2b, are
CC produced by alternative initiation. Both isoforms can activate
CC NF-kappaB. Isoform 1 is more abundant;
CC -!- TISSUE SPECIFICITY: Monocytes-specific.
CC -!- DISEASE: Defects in CARD15 are the cause of Blau syndrome (BS)
CC [MIM:186580], a rare autosomal dominant disorder characterized by
CC early-onset granulomatous arthritis, uveitis and skin rash.
CC -!- DISEASE: Defects in CARD15 are a cause of susceptibility to
CC Crohn's disease (CD) [MIM:266600], a form of remitting
CC inflammatory bowel disease. CD may involve any part of the
CC gastrointestinal tract, but most frequently the terminal ileum and
CC colon. Bowel inflammation is transmural and discontinuous. Crohn's
CC disease is commonly classified as autoimmune disease.
CC -!- DISEASE: Defects in CARD15 are a cause of susceptibility to
CC ulcerative colitis [MIM:191390], a chronic inflammatory bowel
CC disease. In ulcerative colitis, the inflammation is continuous and
CC limited to rectal and colonic mucosal layers. Ulcerative colitis
CC is commonly classified as autoimmune disease.
CC -!- SIMILARITY: Contains 2 CARD domains.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF178930; AAG33677.1; -;
CC EMBL; AF385089; AAK70867.1; -;
CC EMBL; AF385090; AAK70868.1; -;
CC EMBL; AJ303140; CAC42117.1; -;
CC Genew; HGNC:5331; CARD15.
CC MIM; 605956; -;
CC MIM; 186580; -;

DR MIM; 266600; -;
DR MIM; 191390; -;
DR InterPro; IPR001315; CARD.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR007111; NACHT_NTPase.
DR Pfam; PF00560; LRR; 2.
DR SMART; SM00114; CARD; 1.
DR PROSITE; PS0209; CARD; 2.
DR PROSITE; PS0837; NACHT; 1.
KW ATP-binding; Repeat; Leucine-rich repeat; Disease mutation;
KW Alternative initiation; Polymorphism.
FT CHAIN 1 1040
FT CASPASE RECRUITMENT DOMAIN PROTEIN 15,
FT ISOFORM 1.
FT CHAIN 28 1040
FT CASPASE RECRUITMENT DOMAIN PROTEIN 15,
FT ISOFORM 2.
FT INIT MET 28 28
FT DOMAIN 26 122
FT DOMAIN 126 218
FT DOMAIN 293 618
FT NP BIND 299 306
FT REPEAT 786 812
FT REPEAT 814 837
FT REPEAT 842 865
FT REPEAT 926 949
FT REPEAT 954 977
FT REPEAT 982 1005
FT VARIANT 140 140
FT A -> T (in CD and ulcerative colitis).
FT /FTID=VAR_012665.
FT W -> R (in CD).
FT /FTID=VAR_012666.
FT T -> M.
FT /FTID=VAR_012667.
FT R -> C (in CD).
FT /FTID=VAR_012668.
FT L -> R (in CD).
FT /FTID=VAR_012669.
FT P -> S.
FT /FTID=VAR_012670.
FT N -> S.
FT /FTID=VAR_012671.
FT D -> N (in CD).
FT /FTID=VAR_012672.
FT T -> S (in CD).
FT /FTID=VAR_012673.
FT A -> V (in CD).
FT /FTID=VAR_012674.
FT R -> W (in CD and ulcerative colitis).
FT /FTID=VAR_012675.
FT R -> Q (in BS).
FT /FTID=VAR_012676.
FT R -> W (in BS).
FT /FTID=VAR_012677.
FT L -> V (in CD).
FT /FTID=VAR_012678.
FT H -> R (in CD).
FT /FTID=VAR_012679.
FT R -> C (in CD).
FT /FTID=VAR_012680.
FT N -> S (in CD).
FT /FTID=VAR_012681.
FT S -> L (in CD).
FT /FTID=VAR_012682.
FT A -> V (in CD).
FT /FTID=VAR_012683.
FT E -> K (in CD).
FT /FTID=VAR_012684.
FT L -> F (in BS).
FT /FTID=VAR_012685.
FT A -> T (in CD).
FT /FTID=VAR_012686.
FT A -> V (in CD).
FT /FTID=VAR_012687.

FT	VARIANT	684	684	R -> W (in CD).	QY	717	-----ESLHC-----LYETRKNKTLTQV-----MAHPEEMGMCVETDMELLVC	754
FT	VARIANT	702	702	/FTId=VAR_012688.	Db	724	PAAPCEAKSVHAMPQFIWLIRSLYEMQEERLARKAAGLNVGHL-KLTFCSVGPTECAAL	782
FT	VARIANT	703	703	/FTId=VAR_012689.	QY	755	TFCIKFSRHVKKLQI-----IEGRQHRSTWSPTMVVLFRWVPVTDAYWQILFSLVKVT	807
FT	VARIANT	713	713	R -> C (in CD and ulcerative colitis).	Db	783	AFVLQHLRRPVALQLDYNSVGDIGVEQLLPCLVGCKALYLRDNNISDRGICKLIECALHC	842
FT	VARIANT	725	725	R -> C (in CD).	QY	808	RNLKELDLSGNSLSHSAVKSLCKTLRRPRCLLETLRLAGCGLTAEDECKDLAFGLRANOTL	867
FT	VARIANT	725	725	/FTId=VAR_012691.	Db	843	ELOQLALFNNKLTDDGCAHSMAKLLACQNFEL-ALRLGNNYITTAAGAQLAEGLRGNTSL	901
FT	VARIANT	755	755	A -> V (in CD and ulcerative colitis).	QY	868	TELDLSFNVLTDAGAKHLCQRLRQPSCKLQRLQVSCGLTSDCCQDLASVLSASPSELKEL	927
FT	VARIANT	758	758	/FTId=VAR_012693.	Db	902	QFLGFMGNRVGDEGAQAALAEALGDHQ-SLRWLSLVGNNGISVGAQAALAKNVMLEEL	960
FT	VARIANT	778	778	/FTId=VAR_012694.	QY	928	DIQONNLDDVGVRLLCBGL-RHPACKLIRLGDQTT-LSDEMRQELRALEQEKPOLLIIFS	985
FT	VARIANT	793	793	E -> K (in CD).	Db	961	CLEENHLQDEGVCSLAEGCLKKNSSIKLKLSNNCITYLGAEAL--LQALERNDTILEVWL	1018
FT	VARIANT	843	843	/FTId=VAR_012696.	QY	986	R 986	
FT	VARIANT	853	853	N -> S (in CD).	Db	1019	R 1019	
FT	VARIANT	863	863	/FTId=VAR_012698.				
FT	VARIANT	885	885	A -> T (in ulcerative colitis).				
FT	VARIANT	908	908	/FTId=VAR_012700.				
FT	VARIANT	918	918	G -> R (in CD).				
FT	VARIANT	924	924	/FTId=VAR_012701.				
FT	VARIANT	955	955	A -> D (risk factor for CD).				
FT	VARIANT	955	955	G -> D (in CD).				
FT	VARIANT	955	955	V -> I.				
FT	VARIANT	955	955	/FTId=VAR_012703.				
FT	VARIANT	955	955	/FTId=VAR_012704.				
FT	MUTAGEN	305	305	K -> R: NO ACTIVATION.				
FT	SEQUENCE	1040	AA; 115282	MW; 0037592D96D7DDEF CRC64;				
Query Match								
Best Local Similarity				5.2%; Score 393.5; DB 1; Length 1040;				
Matches				198; Conservative 111; Mismatches 325; Indels 147; Gaps 31;				
QY	314	IRDLFG-PGLDQEPRIVILQGAAGIKGKSTLARQVKEAMGRGQLYGDRFQHVFPFSCREL	372					
Db	278	LEELFTPGHLNDADTVLVGEGAGKSTLQRLHLLWAAGQDF-QEFLFVPFSCROL	336					
QY	373	-AQSKVSVLAEL-----TGKGTATPAPIRQILSRPERLLFTLDGVDEPGWVLQEP	422					
Db	337	QCMAPLSVRTLLFEHCCWPVQVEDI-----FQLLDHPDRVLLTFDGFDE--FKFRFT	389					
QY	423	SSELCLHWSQPADALLGSLGKTLTPEASFLITARTTALQNLIPSLQARWEVLGFS	482					
Db	390	DREHCSPTDPTSVQTLNLLQGNLLKWKARVVTSPRAAVSAFLRKYIRTEF-NLKGFS	448					
QY	483	ESSRKEYFYRYFTDERQAIRAFRLVKSNNKELMALCLVPWVSLACTCLMQOMKRKEKTL	542					
Db	449	EQGIELYLRKRHHPEGVADRLIRLLQETSAHGLCHLPVFSMWVSKC-----HQELLL	501					
QY	543	-----TSKTTTTLCL-----HYLAQAL-----QAQPLGP-----QLRDLCSLAAGIW-----	580					
Db	502	QEGGSPKTTTDMYLLILQHLFLHATPPDSASQGLGPSLLRGLRPTLLHLGLRALMGLGMC	561					
QY	581	-----QKXTLFPDDLKXGLDCAIISTFLKMGILQEHPIPLSYSTIHLFCQEFFAA	632					
Db	562	CVVFSAAQLQAAQVSPDDISL-----GFLVRA-----KGVVFGSTAPL--EFLHITFOCFFAA	612					
QY	633	MSYVLEED-----EKGRGKSHSNCIIDLEKTLKLEAYGI	662					
Db	613	FYIALADVPVALLRHLFNCGRPGNSPMARLLPTMCIOASEGKSSVAALLQKA-EPHNL	671					
QY	663	HGLFGASTTRFLGLLSDGEREMENIFHCRLSQGRNL-----MQVPSLQLLLOPHSL-	716					
Db	672	Q-----ITAAFLAGLLSRE---HWGLLABCQTSKALLRRQACARWCLARSRLRKHFSIP	723					

Search completed: January 29, 2004, 13:46:30
Job time : 26.9982 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 13:06:01 ; Search time 79.1934 Seconds
(without alignments)
4656.416 Million cell updates/sec

Title: US-09-996-617-2

Perfect score: 7534

Sequence: 1 MAGGAWRLACYLEFLKXEE.....HLIMELWEKSGKGLPLSS 1429

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phase.*

10: sp_plant.*

11: sp_prodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2640	35.0	509	4 Q9H5Z8	Q9h5z8 homo sapien
2	1754	23.3	409	4 Q9H5Z7	Q9h5z7 homo sapien
3	1371.5	18.2	892	4 Q8NEU4	Q8neu4 homo sapien
4	1254.5	16.7	1033	11 Q8R4B8	Q8r4b8 mus musculus
5	899.5	11.9	982	11 Q8BU40	Q8bu40 mus musculus
6	868	11.5	863	11 Q8C6J9	Q8c6j9 mus musculus
7	843	11.2	846	4 Q8IXT0	Q8ixt0 homo sapien
8	815.5	10.8	713	6 Q9SUZ7	Q9su7 macaca fasc
9	689.5	9.2	748	11 Q99MW0	Q99mw0 mus musculus
10	679.5	9.0	825	11 Q8C6M5	Q8c6m5 mus musculus
11	676	9.0	657	11 Q8BPQ7	Q8bpq7 mus musculus
12	573	7.6	673	11 Q8CCN1	Q8ccn1 mus musculus
13	393.5	5.2	953	4 Q8IWF5	Q8iwf5 homo sapien
14	392	5.2	1020	11 Q8K3Z0	Q8k3z0 mus musculus
15	383	5.1	287	4 Q9BY26	Q9by26 homo sapien
16	372.5	4.9	778	4 Q8NPF48	Q8nfp48 homo sapien

17	363.5	4.8	456	11 Q91V17	Q91vi7 mus musculus
18	361.5	4.8	456	11 Q924P4	Q924p4 mus musculus
19	358	4.8	447	4 Q96FD7	Q96fd7 homo sapien
20	358	4.8	461	4 Q8BQ80	Q8bq80 homo sapien
21	357	4.7	461	6 Q8HZP9	Q8hzp9 pan troglod
22	351.5	4.7	390	11 Q8BWZ1	Q8bwz1 mus musculus
23	351	4.7	953	11 Q8BHB0	Q8bbh0 mus musculus
24	348	4.6	461	4 Q81ZK8	Q81zk8 homo sapien
25	339.5	4.5	516	4 Q8NBF5	Q8nbf5 homo sapien
26	334	4.4	1052	7 Q9GJD8	Q9gjd8 rattus norv
27	334	4.4	1073	7 Q9GJD9	Q9gjd9 rattus norv
28	334	4.4	1153	7 Q9GJE0	Q9gie0 rattus norv
29	325	4.3	269	11 Q9D458	Q9d458 mus musculus
30	318	4.2	1155	7 Q9TPP1	Q9tppl mus musculus
31	316.5	4.2	519	11 Q8C249	Q8c249 mus musculus
32	304	4.0	660	4 Q8NFO6	Q8nfo6 homo sapien
33	291.5	3.9	733	4 Q8TEE2	Q8tee2 homo sapien
34	289	3.8	977	7 Q8HW99	Q8hw99 mus musculus
35	280.5	3.7	884	7 Q29675	Q29675 homo sapien
36	278.5	3.7	932	4 Q96KL4	Q96kl4 homo sapien
37	277	3.7	1056	4 Q8NF42	Q8nfp42 homo sapien
38	276	3.7	692	4 Q96D51	Q96d51 homo sapien
39	272.5	3.6	706	11 Q8BUT6	Q8but6 mus musculus
40	258	3.4	1097	4 Q8H6Y0	Q8h6y0 homo sapien
41	241	3.2	195	6 Q8HXX9	Q8hxx9 bos taurus
42	232.5	3.1	1402	11 Q8CH65	Q8ch65 mus musculus
43	232.5	3.1	1403	11 Q8CH70	Q8ch70 mus musculus
44	226.5	3.0	1402	11 Q8CGT4	Q8cgt4 mus musculus
45	223.5	3.0	1403	11 Q8BG68	Q8bg68 mus musculus

ALIGNMENTS

RESULT 1

Q9H5Z8	PRELIMINARY;	PRT;	509 AA.
ID Q9H5Z8			
AC Q9H5Z8			
DT 01-WAR-2001 (Tremblrel. 16, Created)			
DT 01-WAR-2001 (Tremblrel. 16, Last sequence update)			
DT 01-WAR-2003 (Tremblrel. 23, Last annotation update)			
DE Hypothetical protein FLJ22740.			
OS Homo sapiens (Human).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX NCBI_TaxID=9606;			
RN [1]_TaxID=9606;			
RP SEQUENCE FROM N.A.			
RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,			
RA Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,			
RA Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Ota T., Suzuki Y.,			
RA Obayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,			
RA Isogai T., Sugano S.;			
RT "NEDO human cDNA sequencing project.;"			
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.			
DR EMBL; AK026393; BAB15469.1; ..			
DR HSSP; P10775; 2BNH.			
DR InterPro; IPR001611; LRR.			
DR InterPro; IPR007091; LRR_RNinh.			
DR Pfam; PF00560; LRR; 1.			
DR PROSITE; PS50503; LRR_R1; 2.			
KW Hypothetical protein.			
SQ SEQUENCE 509 AA; 56908 MW; EB7535AF69817F5B CRC64;			

Query Match 35.0%; Score 2640; DB 4; Length 509;
Best Local Similarity 100.0%; Pred. No. 2.2e-190;
Matches 506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	735	MAHFBEMGMCVETDMLLVCTCFIKFSRHVKQLIEGRQHSTWSPTMVVLFWRVPVTD	794
Db	1	MAHFBEMGMCVETDMLLVCTCFIKFSRHVKQLIEGRQHSTWSPTMVVLFWRVPVTD	60
Qy	795	AYWQILFVLKVTNLKELDLGNSLSHSAVSKLCKTLRRPRCLLETTLRAGGLTAEDC	854

Db 245 DYLFYINCKMNSQATECSMDQLIFSCWPEPSAPLQELIRVPERLLFIIDGDELKPSPH 304
QY 421 EPSSELCLHWSQOPADALISLGLKTIILPEASFLITARTTALQNLIPLEQARVVEVLG 480
Db 305 DQGPWCLWEERPTPELLNSLRKLPESLLITRTPTALEKLRHLEPRHVEILG 364
QY 481 FRESRKEYFYFTDERQAIIRAFRLVKNKELMALCLVPWVSWLAECTCIMOQMKRKEKL 540
Db 365 FSEARKEYFYFHNAEGOVFNVRDNEPLFTWCFVPLVWVWVCTCLOQLEGGGILL 424
QY 541 TLTSKTTTTLCHLYLAQAQPLGPOL-----RDCLSLAAEGWQKTLFSPDRLKH 594
Db 425 RQTSRTTAVMDYLSLUSLQKPGAPRLQPPNQRGLCSLAADAGMNQKILFEEQDLRKH 484
QY 595 GLDGALISTFLKMGILQEH-PIPLSYFHLFCQEFFFAAMSYYL-EDEKRGKHSNCIID 652
Db 485 GLDGEDVSFAFLNNIFQKDINCERYYSFHLSPQEFFFAAMYILDEGEGAGPDQ----D 540
QY 653 LEKTLBAYGI-HGLFGASTTRFLGLLSDEGREMENI FHCRLSQ--GRNLMQWVPSLQ- 708
Db 541 VTRLLTEYAFSERSFLATSRFLGLNBEETRSHLEKSLCWKVSPIHKMDLLQWIQSKAQ 600
QY 709 ---LLIOPHSLESCHLYETRNKTLFTQVMAHFEEMGMC-VETDMELLVCTFCIKESRHV 764
Db 601 SDGSTLQOGSLEFFSCLYEQIEEFTQALSHFQVIWNSIASKWBHMYSSFCCLKRCSRA 660
QY 765 KKLQILT-----EGRQHRSTWSPWVFLFR--WVPVTDAYWOILFSLVKVTRNLKELD 814
Db 661 QVHLHYGATYSADGEDARCSAGAHTLLVQLRPRTVLIDAYSEHLAAALCTNPNIELS 720
QY 815 LSGNSLSHSAVKSCLKTLRRPRCLLETLRLAGCLTAEDCKDLAFLGLRANQTLTDLDSF 874
Db 721 LYRNALSGVGLLCCQLRHPCNKLQNLRLKRCRISSACEDLSAALANKNLTRMDLSG 780
QY 875 NVLTDAGAKHLCORLPQPSCKLQRLQVSGCLTSCDCCDLASVLSASPSEKELDQONNL 934
Db 781 NGVGFPGMMLCEGLSHPCRLQMIQLRKQLESAGACQEMASVLTGNPHVELDTGNAL 840
QY 935 DVGVRLLCEGLRHPCAKLRL-----GLQDTLSDEMROELRALEQEKPOL 981
Db 841 EDLGLRLCCQLRHPCRLTLWLFGMD---LNMKTHSRLAALRVTKPYL 888

RESULT 4
ID Q8R4B8 PRELIMINARY; PRT; 1033 AA.
AC Q8R4B8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Mast cell maturation inducible protein 1.
GN MWIG1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/cJ;
RA Kikuchi-Yanoshita R., Koga K., Taketomi Y., Sugiki T., Saito T.,
RA Ishii S., Hisada M., Suzuki-Nishimura T., Uchida M.K., Moon T.-C.,
RA Chang H.-W., Sawada M., Inagaki N., Nagai H., Murakami M., Kudo I.,
RT Identification of inducible genes during in vitro maturation of mouse
RT bone marrow-derived mast cells to connective tissue-type mast cells.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF486632; AAL90874.1;
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR001611; LRR_
DR InterPro; IPR007091; LRR_Rninh.
DR InterPro; IPR003590; LRR_Rninh_sub.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF00560; LRR; 2.

DR Pfam; PF02758; PAAD_DAPIN; 1.
DR SMART; SM00368; LRR_RI; 1.
DR PROSITE; PS00824; DAPIN; 1.
DR PROSITE; PS00038; HLH 1; 1.
DR PROSITE; PS00503; LRR_RI; 3.
DR PROSITE; PS00837; NACHT; 1.
SQ SEQUENCE 1033 AA; 118274 MW; 5924690966B12117 CRC64;
Query Match 16.7%; Score 1254.5; DB 11; Length 1033;
Best Local Similarity 32.6%; Pred. No. 2.2e-85;
Matches 331; Conservative 164; Mismatches 334; Indels 187; Gaps 23;
QY 8 RIACYLEFLKKELEKEPOLLANKAHSSSSSETPAQPEKTSKMEVASYLVIAQYGEQRAW 67
Db 7 KLAQYLEDLEDVLLKKFKMHLEDYPPKGCIPVPRGOMEKADHLDTATIMDFNGEEKAW 66
QY 68 DLALHTWQMGLRSICAAQEGAGHSPSPFPYSPSEHLGSPSQPTSTAVLMPWIELPAG 127
Db 67 ANAVMIFAANRRDLWEKAKK-----DQP-----EWNDT 95
QY 128 CTQGSERRVRLQPTSGRRWRREISASLLYQALPSPDHESPSQSPNAPTSTAVILSGWG 187
Db 96 CTSHSS-WVQSE--DSLSEW-----MGLLG 118
QY 188 SPQPSLAPREQAPGTOWPLDETSGIYYEIREREREKSEKGRPPMAVWGTTPQAHTS 247
Db 119 YLSRISICKKK-----DYCKMYRRHVRSRFYSIKDRN-----ARLG----- 155
QY 248 LQPHHPWEPVSRESLCSTWPKNEDFNQKFTQLLLQRPHPRSQDPLVKRSPDYVEEN 307
Db 156 -----ESVDLSRYTQLQV-KEHPSKOE----- 178
QY 308 RGHLLI-----EIRD-----LFGPLDTQEP-RIVILQGAAGIGKSTLAROVKEA 350
Db 179 REHELLTTRTKWRDSDPMSSLKLELLFEDCHSEPVHTVFGAAGIGKTLARKIMLD 238
QY 351 WGRGOLYGDROFQHYFYSFSCRELAQSKVVSALBELIKGOGTATPAPIRQILSRPERLLFILD 410
Db 239 WALGKLFKDKDYLFFIHCREVSLRTPRSLADLVSCWPDNPVPVKILRKPSRILFLMD 298
QY 411 GVDEFGWLQVSPSELCHWSQOPADALLGSLGKTLILPEASFLITARTTALQNLIPSL 470
Db 299 GFDELQGADEFHIGEVCTDWQKAVRGDILLSSLRKLLPKASLLITTPVALEKLUHL 358
QY 471 EQARVVEVLGPFSESRKEYFYFTDERQAIIRAFRLVKNKELMALCLVPWVSWLAECTCL 530
Db 359 DHPRIVEILGPFSEAKRKEYFFKFNELQARAFRLIQENEVFTWCFIPLVCNIVCTGL 418
QY 531 MQQMKRKEKLTLSKTTTTLCHLYLAQAQAP-----LGQRLDCLSLAAEGIWQKKT 584
Db 419 KQOMETGKSLAQTSTKTTTAVVVFLLSLQSRGGIEEHLFSDYLGCLCSLAADGWNQKI 478
QY 585 LPSDDLKHKHLDGAIISTFLKMGILQ-EHPIPLSYFHLFCQEFFFAAMSYYLDE---- 640
Db 479 LPEECDLARKHGLQKTDVSAFLRMNVQKEDCERFYSFSHMTFQBFFAAMYLLLEEARG 538
QY 641 ----KGRGKHSNCI-IDLEKTLQAYG-IHGLFGASTTRFLGLLSDEGREMENIFHCRL 694
Db 539 ETVRKGPGGCSDLLNRDVRKLLNFKYKFKGKGVLI FVVRFLFGLVNOERTSYLEKLSCKKI 598
QY 695 SQ--GRNLMQWV-----PSLQLLQPLSHLESCHLYETRNKTLFTQVMAHFEEMGMCVETD 748
Db 599 SQOVRLELLKWLIEVKAKAKKQWQSQLELFCYLEMQEEDFVQSAMDHFFKIEINLSTR 658
QY 749 MELLVCTFCIKFSRHVKKLQ-----IEGRQHRSTWSP-----MVVLFRWV 790
Db 659 MDHVVSFCIKNCHRVKTLISLGFHNSPKEEERRRGGRLDQVQCVFPDTHVACSSRLV 718
QY 791 P--VTDAYWOILFSLVKVTRNLKELDLSGNSLSHSAVKSCLKTLRRPRCLILETLRAGCG 848
Db 719 NCLTSSFCRGLFSSLTNRSLTLEDNTLGDPMRVLCBALQHPGNCIQRMLGRGC 778
QY 849 LTAEDCKDLAFLGLRANQTLTDLDSFNVLTDAGAKHLQRLRQPSCKLQRLQVSGCLTS 908

```
Db 779 LSHQCFFDISSVSSSQKVELDSDNALGDFGIRLLCVGLKHLNCLQKMLVSCCLTS 838
Qy 909 DCCODLASVLSASPSPSLKELDLQONNLDVGVRLLCGLRHACKLIRLGLDQTTLS 964
Db 839 ACCQDLALVLSNHSUTRIYIGENALGDSGVQVLCCKMKDPQCNLQKGLVNSGLT 894

RESULT 5
Q8BU40 PRELIMINARY; PRT; 982 AA.
AC Q8BU40;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Similar to PAN2 protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK087843; BAC40024.1; -.
SQ SEQUENCE 982 AA; 112600 MW; 6E2DBCFEA2053585 CRC64;

Query Match 11.9%; Score 899.5; DB 11; Length 982;
Best Local Similarity 26.8%; Pred. No. 1.2e-58;
Matches 269; Conservative 176; Mismatches 366; Indels 191; Gaps 26;

Qy 203 GTQWPLDETSGIYYTIRE-REREKEKG--RPPMAAVVGTPTPQAHTSLQPHH----PW 255
Db 9 GLMWYLELNKKEFVKFKFLQEVQLQGLKQVSWTEVKKASRDOLASLLKHYEKPAM 68
Qy 256 EPSVRESLCTWPKWKNEDNQKFTQLLLORPH-----PRSQDPLVKRSPDYVEENRG 309
Db 69 DMTFR-----EFQIKNRKDLIERAKREIDGCPKLYRAHMKTKMTH--DSSRA 113
Qy 310 HLIEIR-----DLQPGLDTOEPRIVILQGAIGKGTATPAPIRQILSRPERLLFILDGVD 353
Db 114 FTISQNFLEKFTEDDYCFENLFOSKGTESKQPVFLSGGAGVGTLMKRLMLAWIE 173
Qy 354 GOLYGRFORHVFVFSRELAQSKVSLAELIGKGTATPAPIRQILSRPERLLFILDGVD 413
Db 174 SPVFLHKFSYIYFCREVKQLKTASLAEILSREWPGPSAPIEELISKEKLLFIIDSL 233
Qy 414 EPGWVLPESSELCHWSQPADALLGSLGKLTILPEASFLITARTALQNLIPSLEQA 473
Db 234 GMECDLFKWESELCDNCTEKQPVNLLSLLRRKMLPESLLISATPESFEKWNRIEY 293
Qy 474 RWVEVLGFSESRKEFYFYFDTERQAIAPRFLVSKNKLWALCLVPWVSWLACTCLMQQ 533
Db 294 HVKIITKGLKERNIKSFHFLFDNRNHEAFSLVRENEQLFTVCQPVLCVWVATCLKEE 353
Qy 534 M-KRKEKLTSTKTTTTLCHVLAQALQALPGP-----QLRDLCSLAAEGWQKKTFL 586
Db 354 IEKGRDPVICRCTSLYTHLNFENFIPQNAHSPSKSQDQLOGLCSLAEGWMTDTFFV 413
Qy 587 SPDDLKHLGDGAIIITFLMGIL-QEHPILPSYFIHLCFQEFFAAMSIVLEDEKGRK 645
Db 414 GEEALRRNGIMSDIPILLDIGMLINIRESEKSYIFLHPSVQEVCAAIFFYLL--KSHVD 470
Qy 646 H-SNCIIDLEKLEAY-----GIHGLFCATTFFLIGLLSDEGEREMENIFHCLRSQ--GR 698
Db 471 HPSQEVKSIKLMFAFLKVKVQWIFFGS---FIFGLLHSESKQKLEAFHGQLSQETIK 527
Qy 699 NLMQWVPSL-----QLLLQPSLESLSHCLYETRNKFTLTQVMAHFEBMGVCVETMELLVC 754
```

```
Db 528 QLYQCLETISGNEELQEQIDGMKLFYCLFEMDDDTFLVEAMNCEQINFVAKDYSDVIVA 587
Qy 755 TFCIKFSRHVKL-----QLIEGRQHS-----TWSPTWVL-----FRWVPV 792
Db 588 AHCLKHCTFLKKLSFSTQNVLSGAQHSYMERLLLTAMNHICSVFIISKDIOELRMKDTNL 647
Qy 793 TDAYWQIILFSLVKVTR-----NLKELDLSGNSLSHSAV 825
Db 648 SESAFSVLYNNLKYHNYTLNVLVANNVFFVCEKYLPFELIQNCNLQHLNLSITILSHSDV 707
Qy 826 KSLCKTLRRPCLLETTLRACGLTAEDCKDLAFGLRANQTLTDLDFNLVTLTDAGAKHL 885
Db 708 KLLCDVLSQAECNTEELVVAACSISSDDCKVFASVLISNKTLLKHLNLSNTL-DKGIASL 766
Qy 886 CORLRQPSCKLQRLQVSCGLTSDC-----NLDYNGLVWLCE 939
Db 767 CKSLCHPDCILHLVLANCSLNEKCDVYLSEVLRNKTLSHLDISSNLDKDBGLKVLCA 826
Qy 911 -----CODLASVLSASPSLKELDLQONNLDVGVRLLCGLRH 948
Db 827 LTLPSVLKSLSLRHCLITISGCQDLAEVLNNQNLVLSQVSNKLEDTGVKLLCDAIKH 886
Qy 949 PACKLIRLGLDQTTLSDEMROELRALEQKPOLIFSRKPSVMTPTTGLDTGEMSNSTS 1008
Db 887 PNCHLEDLGLACELTGACCEDLASTFTQCKTLWAMNLLK-----NALDYNGLVWLCE 939
Qy 1009 SLKQRLGSERAASHVAQANLKLDV-SKIFPIAIEAESPP 1049
Db 940 ALKQO-----TCATYVLGLQITDFDTETQAFILVAE--QEKNP 974

RESULT 6
Q8C6J9 PRELIMINARY; PRT; 863 AA.
AC Q8C6J9;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical RNI-like structure containing protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK054426; BAC35775.1; -.
KW Hypothetical protein.
SQ SEQUENCE 863 AA; 99955 MW; DJA46C782DA544CF CRC64;

Query Match 11.5%; Score 868; DB 11; Length 863;
Best Local Similarity 27.7%; Pred. No. 2.4e-56;
Matches 236; Conservative 154; Mismatches 343; Indels 120; Gaps 17;

Qy 214 IYTYIERERERE-----KSEKGRPPWAA-----VVG---TPQANTSLQPHHPWP 257
Db 41 ISWTEKASREDLAILLVKHCQDQAWDTTFRVQMGIRNVITNRAIGETAHSTIYRA 100
Qy 258 SVRESLCTWPKQEDFNQKFTQLLLQRPSPQDPLVKRSPDYVEENRGHLIEIDL 317
Db 101 HLKEKL-----THDCSRKE-----NISQNFQDEYDHL---ENL 132
Qy 318 FPGPLDQEPRIVLQGAAGIKSTLARQVKEAMRGOLYGRFOHVFYFSCRELAQSKV 377
Db 133 LVPNGTENNPRKVVVLQGVAGIGKTIKLLKMLVMSEGLVFQNKFSYIFFCCHDKQLQT 192
```


QY 378 VSLAEILGKGTATPAPIRQILSRPRLFLILGVDVPCGWVQLQEPSSSCLHWSQPQPAD 437
Db 193 ASLADLISREWPSPSPAMEILLISQPEKLLFIIDSLEGMEWNTQDSQLCYNCKMEQPVN 252
QY 438 ALGSLGKLTILPEASFLITARTALONLIPSLQARWVEVLGFSSESSKEYFYRYFTDE 497
Db 253 VLLSSLRKILPESSLLSTSCETPKOLKOWIEYTNVTTIGFKENNINMCFHSLFQDR 312
QY 498 ROAIRAFRLVGNKELWALCLVPWVSWLACTCLMQOMKR-KEKLTILTSKTTTCLHLYLA 556
Db 313 NIAQEAFLIRENEQLFTVQAPVVCYMWATCLKNEIESKOPVSICRTTSLTYTHILN 372
QY 557 QALQAOPLGQ-----LRDLCSLAEGTWKKTFLSPDDLRKHGDLGDAIISTFLKMGIL 610
Db 373 LFTPHNAQNPNNSDLDNLCLFVAVEGHWTDISVNEEALRRNGIMDSIPTLLDIGIL 432
QY 611 -QEHPIPLVSFTHLCFOFFAAMSVLDEKGRGHSNCID-----LEKTLAYGHIH 663
Db 433 EOSRESSENSYIFHPVSQVBFCAAMFYLLHSEMDHSQGVYFIETFLTFLNKKIKQWFL 492
QY 664 GLFGASTTFLGLLDEGEREMENIFHCRLSQ--GRNLMQWVPSLQLLQP-----HSL 716
Db 493 GC-----FFGLLHETQEKLEAFGYHLSKELRRQLFLMLELLDTHLPDVKKINTM 545
QY 717 ESHLCYETRNKTLTQVMAHFEEMGCVETDMELLVCTFCIKF-----SRHVKK 766
Db 546 KFFCYCLFEMEEVFOVSAMNCRQIDVWVKYSDFVAAVCLSHGSALTDFSIQAQNVLN 605
QY 767 LQIEGRQHRSTWPTMVVLFWRVPV-----TDAYWQLFSVLK----- 805
Db 606 BELGORGKLLILHQCISVFLRNKDKITURIEDTIFNEPVKIFYSYLKNSCILKTLVA 665
QY 806 -----VTNKLKELDSGNSLSHSAVSKLCKTLRRPRCLLETLRLAGCGL 849
Db 666 YNVSLFCDKRLFLLELIQSINLELYLRTGLSHSDVEMLCDILNQECNIRILDANCSL 725
QY 850 TABDCKDLAFGLRANQTLTDLDSFNVLTDAKAKHLQRLRQPSCKLQRLVLSQGLTSD 909
Db 726 CEHSWYLSVLRQNSLRNYNISYNNLKDGLKALCRALTLPNSALHSLSLEACQLTGA 785
QY 910 CCODLASVLSASPSLKELDLQNNLDDVGRLLCEGRHPACKLIRGLDDOTTLSEMRQ 969
Db 786 CCKDLASTFTRYKCLRINLRNKLAKNSGSLFVLCKAMKDQCTLTYELKLRMADF-DSDSQ 844
QY 970 ELRALBOEKPOL 982
Db 845 EFLSEMERKIL 857

RESULT 7

Q8IXT0 PRELIMINARY; PRT; 846 AA.
AC Q8IXT0; 2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DE Similar to NALP2 protein.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strausberg R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC039269; AAH39269.1; -
SQ SEQUENCE 846 AA; 96368 MW; 46BB8245550E39F7 CRC64;

Query Match 11.2%; Score 843; DB 4; Length 846;
Best Local Similarity 30.0%; Pred. No. 1.8e-54;
Matches 255; Conservative 154; Mismatches 328; Indels 114; Gaps 28;

QY 148 WREISASLYQALPSSPDHESQSPNAPTSTAVLGSWGSPPQPSLAPRQEAFTQWP 207
Db 67 WYEMASLOVFEKM-----HRMDLSBRAKDEVREAALKSFNKRKPLSLGTTIKERP-----P 117
QY 208 LDETSIYYTELREREREKSEKGRPWAAVVGTPPQAHTSLQPH-----HHPWESVRES 262
Db 118 LD-----VDEMLERFKTEAQAFETTKGNVICLGEVFGKKPKDKNRCRYILTKTKFRE- 170
QY 263 LCSTWPKNEDFN---QKFTQLLLQRPHPRSQDPLVKRSWPDYVEENRGLHIEIRDFG 319
Db 171 MKSWPGSDSKVQWMAERYKMLIPSNP-----RVLP 203
QY 320 PGLDTQEPRIVLGAGAGIGKSTLARQWKEAGRQOLYGRPQHVYFVSCRELAOSKVVS 379
Db 204 PFSYT-----VVLGYPAGLGKTTLAQKMLDWAEDNLI-HKPKYAFYLSCRELSLGPSC 257
QY 380 LAELIGKGTATPAPIRQILSRPRLFLILGVDVPCGWVQLQEPSSSCLHWSQPQPAD 437
Db 258 FAELVFRDWEPELQDDIPHILAQARKILFVIDGDELGAAPGALIE-----DICGDEWKKAP 313
QY 436 ADALGSLGKLTILPEASFLITARTALONLIPSLQARWVEVLGFSSESSKEYFYRYFT 495
Db 314 VPVLGSLNLRVMLPKAALLVTRPRALDRILAEPIYIRVEGFELEEDRAYFLRHFG 373
QY 496 DERQAIRAFRLVGNKELWALCLVPWVSWLACTCLMQOMKRKEKLTILTSKTTTCLHLYL 555
Db 374 DEDQAMRAFELMRSNAALFQLGSAVAVCWIVCTTLKQMEKGEDPVPTCLTRTGLFLREL 433
QY 556 AQAL-QAOPGLGQRLDCLSLAEGTWKKTFLSPDDLRKHGDLGDAIISTFLKMGILQEH 614
Db 434 CSRFQGAQLRGALRTLSLAAQGLWAQTSVLHREDLERLGVQESDLRLFLDGLDILQDR 493
QY 615 IPLS-YSFTHLCFOFFAAMSVL--EDEKGRGHSNCIIIDLEKTLQYAGIHLFGASTT 671
Db 494 VSKGYSYFIHLSFQFLTALFYLEKEBEDRDGHTWIDGVQKLS--GVERLNPDLI 551
QY 672 R---FLGLLSDGEGREMENTIFHCRLSQGRNLMQWVPSL-QLLQ-----PHSL---- 716
Db 552 QAGYYSFGLANEKRAKELEATEFGCRMS-----PDIKQELLRCIDISCKGHSVTVDL 602
QY 717 -ESLHCLYETRNKTLTQVMAHFEEMGCVETDMELLVCTFCIKFSRHVK--LQIL-BG 772
Db 603 QELLGCLYESQEEELVKEVMAQFKEISHLNA-VDVVPSSFVCKVGRNLRQKSLQVIKEN 661
QY 773 RQHRSTWPTMVVLF-----WVPVTDAYWQLFSVLKVRNKLKELDSGNSLSHSAV 826
Db 662 LPENTVASESDAEVRSQDDQHMLP-----FWTDLCSI FGSNKDLMGLAINDSFLSASLVR 717
QY 827 SLCKTLRRPRCLLETLRLAGCGLTAEDC-KDLAFLGRANQTLTDLDSFNVLTDAKAKHL 885
Db 718 ILCEQIASDTCHLQ--RVVFKNISPADAHRLCLALRGHKTVTLTLOGNDQDDM-FPAL 774
QY 886 CORLRQPSCKLQRLQVSCGLTSDCCODLASVLSASPSLKELDLQNNLDDVGRLLCEG 945
Db 775 CVELRHPECNRYLGLVSCSATTQOWADLSLAEVNSQSLTCVNLSDNELLDEGAKLLYTT 834
QY 946 LRHPACKLIRL 956
Db 835 LRHPCKFLQRL 845

RESULT 8

Q95LZ7 PRELIMINARY; PRT; 713 AA.
AC Q95LZ7;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical 80.3 kDa protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.

```

OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Hashimoto K., Osada N., Hida M., Kusuda J., Tanuma R., Hirai M.,
RA Terao K., Sugano S.;
RT "Isolation of novel full-length cDNA clones from macaque testis cDNA
RT libraries";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB071044; BAB64437.1; --
DR InterPro; IPR007091; LRR RNinh.
DR DR InterPro; IPR007111; NACHT NTPase.
DR PROSITE; PS0503; LRR RI; 2.
DR PROSITE; PS0837; NACHT; 1.
KW Hypothetical protein
SQ SEQUENCE 713 AA; 80258 MW; 6F214C9B773F54DC CRC64;

Query Match 10.8%; Score 815.5; DB 6; Length 713;
Best Local Similarity 31.4%; Pred. No. 1.6e-52;
Matches 228; Conservative 114; Mismatches 248; Indels 135; Gaps 18;

QY 383 LICKDGTATPAPRIQLSRPERLLFILDGVE----PGWVLPSPSELCLHWSQOPADA 438
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 MVRFDWPELQDDIPHLIAQAQKILFVIDGFELGAPPGALIQ----DICGDWEQKPEVPV 56
QY 439 LIGSLGKTLIPASLTARTTALONLIPSLQARWVEVLPFSSRSKEYFYRYFTDER 498
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
57 LIGSLKRMKLPATLLVTRPALRDRLFLAQPDIYIRVEGFLEDRRAYFLRHFGDDE 116
QY 499 QAIRAFRLVSKNKLWALCLVPVWSWLACTCLMQMKREKLTLSKTTTTLCLHYL-AQ 557
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
117 QAMRAFELMRSNAALFQLGAPAVCVICTLKLQMEKGEDPAPTCLTSTGLFLRFLCSQ 176
QY 558 ALQAQPLGQRLDCLSLAAGIHWQKTLFSPDDLRKHLGDGAIITFLKMGILQHEPIPL 617
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
177 FPQGAQLGALRALSLAASLQAQSLWAQMSVLHGEDLESAGVQESDLRLFLDGLILRODGVAK 236
QY 618 S-YSFTHLCFOEPFAAMSYYL--EDSKRGKHSNCIIIDLEKTL--EAYGIHLGFASTT 671
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
237 GCYSFHLHFSQQLTALFVALEKEEEDRDGAWDLDGVQKLSREERLKNFDLQAG-- 294
QY 672 RFLGLLSDEGEREMENIFHCLRSQGNLWQWVPSL-QLLLQ-----PHSLS 718
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
295 RFLFGLANERKVKLEATFGWRMS-----PEIKQELLRCDSVRKNGHFTAADLREL 345
QY 719 LHCLYETRNKTLFTQMAHFEEMGMVETDMELLVCTFCIKSRHV-----KKLQ----- 769
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
346 LCCLYESQEDLVKEYMAQFKEISHLNA-VDIAPSSFCFKHCQNIFFADACRNLCALR 404
QY 770 -----IEGRQHRSTWSPTMVVL-----PRWV-----PVTDAYWQLFSVLKVTEN 809
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
405 GHKVTHTLTLQGDQDMLPALCEVLRHPECNRYLGLVSCSNTTOQWADLSLAEANRS 464
QY 810 LKELDLSGNSLSHSAVKSCKTLRRPRCLLETLRLAGCGLTADCKDLAFGLRANQTLTE 869
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
465 LMCVNLSDNELLDEGAKLYTLRLHFKCFQLRSLNCHLTEANCCKDLAAVLVWSRELTH 524
QY 870 LLSFNVLTDAGAKHICQRLRQPSCKLQRLQVSCGLTSDCCODLASVLSAPSCLKELD 929
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
525 LCLAKNSLXDTGVKFLCEGLSPECKLQALVWNCDDITSDGCCSLAKLQEKSSLSCLDL 584
QY 930 QQNLLDDVGRLLCEGLRHPACKL----- 953
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
585 GLNHIGVTGVKVLCEALSPLCNLRCNLWGCISPPFSCEDLCSALSCNQSLTLDLQGN 644
QY 954 -----IRLGLDQTTLSDEMQLRALEQKPOLIFPS----- 985
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
645 PLGSSGVKMLFKTLTRPGTLQTLRLKIDD--FNDELHKLLEEENNPQLIIDTEKHDPW 702
QY 986 RKRPS 990
DB : : : :
703 KKRPS 707
```

```

RESULT 9
Q99MW0 PRELIMINARY; PRT; 748 AA.
AC Q99MW0;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Ribonuclease/angiotensin inhibitor 2.
GN RNH2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=21175748; PubMed=11279525;
RA Wang P.J., McCarrey J.R., Yang F., Page D.C.;
RT "An abundance of X-linked genes expressed in spermatogonia.";
RL Nat. Genet. 27:422-426(2001).
DR EMBL; AF285581; AAK31960.1; --
DR HSSP; F10775; 2BNH.
DR MGD; MGI:1890518; Rnh2.
DR InterPro; IPR007091; LRR RNinh.
DR PROSITE; PS0503; LRR RI; 1.
SQ SEQUENCE 748 AA; 84946 MW; 674A385E011DE8EC CRC64;

Query Match 9.2%; Score 689.5; DB 11; Length 748;
Best Local Similarity 26.2%; Pred. No. 5.6e-43;
Matches 197; Conservative 120; Mismatches 241; Indels 193; Gaps 14;

QY 419 LQEPSELCLHWSQOPADALLGSLGKTLIPASFLITARTTALONLIPSLQARWVEV 478
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
5 MSERESELCDCTEKQPLRILLSLRRMLPKSFLISATPETERKEGVECTNNKIV 64
QY 479 LCFSSRSKEYFYRYFTDERQAIRAFRLVSKNKLWALCLVPVWSWLACTCLMQQM--RRK 537
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
65 TGFNESNIKMVFRSLFQDKTKTQEIFSLVKNQQLFTVCQVPLCVMTATCLKKEIKGR 124
QY 538 EKLUTSTTTTTLCLH-----YLAQALQ--AQLGQPLRDCLSLAAGIHWQKTLFSPDDL 591
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
125 DLVSVCRRTTSLYTHIFNLFIQSAQYPSKESQAQLQSLCSLAEGMWTDTTFVFEAL 184
QY 592 RKHGLDGAISITFLKMGILQE-HPIPLSYTHLCFOEFFAAMS YVLEDEGRGHSNCI 650
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
185 RRMGIMSDIPITLDDVRILEKSKKSEKSYIFLHPSIQVCAAFYLL---KSHMDHPQSD 241
QY 651 IDLEKTLQA-----YGIHGLFGASTTRFLLGLLSDEGEREMENIFHCLRSQ--GRNL 700
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
242 V---KSIEALIFTFLKKYKQVWIFPGS---FIFGLHSEKQKLEAFFGHQLSQEIKRQL 295
QY 701 MQWVPSL-----QLLQPHSLSLHCLYETRNKTLFTQMAHFEEMGMVETDMELLVCTF 756
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
296 YQLETISGNELOQVDGMKLFYCLFEMDDAEFLAQMCMQEQINFVAKDYSDVIVAHA 355
QY 757 CIKFSRHVKLQ-----LIEGRQHRST-----W----- 779
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
356 CLQHCSTLKKLSLSTQNVLSEGEHSYTEKILMCMHHMCSVLISSKOYIILQVKNLNE 415
QY 780 ----- 779
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
416 TASLVLYSHLWVPSCTLKAALVNNVTVFLCDNRLFPFELIQNCLQHLDLNLTLFSLHGDVKL 475
QY 780 -----SPTMVVLFRWV----- 790
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
476 LCDVLSQBECEKIEKLWAAACNLSPDCKVFASVLISSRMLKHLNLSNNLDKGISLSKA 535
QY 791 -----PVTDAYWQLFVSLVTRNLKELDLSGNSLSHSAVKSCKLCTLR 834
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
536 LCHPDCVLKNLVLNCSLSEQCQWDYLSVLRNRKTNLHLDISSNDLKDEGLKVLCLRALSL 595
QY 835 PRCLLETLRLAGCGLTADCKDLAFGLRANQTLTDLFSLFNVLTDAGAKHLQORLPSC 894
```

```

Db 596 PDSVLKSVRYCLITSGCQDLAEVLRKNQNLRLQVSNKIEDAGVKLLCDAIKHNC 655
QY 895 KLORLQVSCGLTSCCCODLAVLGSASPSLKELDLQNNLDVGVRLLCGLRHHPACKLI 954
Db 656 HLENGLEACALTAGCCEDLASAFHCKTLWGINLQENALDHSGLIVLFEALQKQOCTLH 715
QY 955 RIGLDQTLSDMRQELRALEQKPOLIFS 985
Db 716 VLGLRITDFDKETQELLMAEEKNHLSLS 746

RESULT 10
Q8CGM5 PRELIMINARY; PRT; 825 AA.
AC Q8CGM5;
DT 01-WAR-2003 (TREMELrel. 23, Created)
DT 01-WAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-WAR-2003 (TREMELrel. 23, Last annotation update)
DE Weakly similar to PAN1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium.
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL: AK054264; BAC35710.1; -.
SQ SEQUENCE 825 AA; 94021 MW; 11D71DEABEAFCA31 CRC64;

Query Match 9.0%; Score 679.5; DB 11; Length 825;
Best Local Similarity 26.0%; Pred. No. 3.7e-42;
Matches 212; Conservative 150; Mismatches 313; Indels 141; Gaps 24;

QY 208 LDSTGIYVTEIREREREKSEKRPWAAVGTPQAHSTLQPHHPWEPVSRLCSTW 267
Db 76 LEEINRVDLAELVQKIEB-----AVLKVPKEKVS-----KPREPSG-----TLTF 117
QY 268 PWKNEDFNQKFTQLLLQPHRPSODPLVKR-----SWPD-----YVENRGHLIE 313
Db 118 PW-----NFVQ--CAKRPEDQKKEWTRTAKQNFWKPKCKNKEIYVVTESKTLIA 168
QY 314 IRDLFGPLDTPQPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGDRFOHFVYFSCRELA 373
Db 169 ---LCNPKIETPFAHAILVHGGPGSGKTTWAKQMLEWSESK-QAQIFSCAFYISCREVN 224
QY 374 QSKVSVLAEIKGDTATPAPIRQILSRPERLLFILDGVDEPGVWLQEPSSSELCHWSQP 433
Db 225 NTKPCTTFAHLSMDNPSWDCVIRDILGKEFLFVVDGDELTFPAGALIRDLCDGDNVT 284
QY 434 QPADALLGSLGKTLTPEASFLITARTALONLIPSEARVVEVIFGSESSKEYFYR 493
Db 285 KPEVLLGSLKRNKMPHATLVITRTQSLHQIFVMDQDPLIVETLGLFLEQKQEFYQY 344
QY 494 FTDER-----QAIRAFRLVKSNNKELWALCLVPWVSWLACTCLMQMKRKEKLTITS 544
Db 345 FEDEGEEDKGEKALRAKALKEVRCNADLYQMASLPTACGIFCLCLERMKKGEDLSLTC 404
QY 545 KTTTLCILHYLAQAQLPGPOLRD-----LCSLAEGIWQKTLFSPDDLRKHGL 596
Db 405 QYTSMLNFLCEVFSSETCEDHLNEEFQILFKKICILAANSLLEQVPLCBEDF----- 459
QY 597 DGAIISTFLKMGILQEHPIPLSY-----SFIHLCEQEPFAANSVYLE---DEK 641
Db 460 -----LTLKLNLNHPVWCRHILFKDSSSTHCSUSFICLGIOQLLAAIIFVOELGQESK 513
QY 642 GRGKHS-NCIIIDLEKTLVAYGIHGLFGASTTRFLGLLSDEGEREMENIFHCRLS----- 695

```

```

Db 514 GVSKYIQNNMLSRBARLKNPDLGSL-----PFVGLLNTRIQLKTKTFGCQISTEVKR 568
QY 696 -----QGRNLNQWVPSLQLLQPHLSLESRLHCLVETRKNKTLFTQVMAHFEENGMCVETDM 749
Db 569 KFLCESEEN-----KPLLLNMNQEILSCLYESQEGFVKEAMVLFDISLHLKST 621
QY 750 ELLVCTFCIKFSRHVKLQL-IE-----GRQHRSTWSPMTWVLFWRVVPVD 794
Db 622 DLIHASFCLKNQNLQTMSLKVERAVFPENVAALSTAKHQRSPDEQRL----- 671
QY 795 AYWQILFSLVKVTRNLKELDLSGNSLSHSAVSKLCTLRPRCLLETLRACGLTAEDC 854
Db 672 TFWTDFCDTFNSNKKLVFELDTHEFLNSSALEILCEKLPSACCCKQKVLK--NISPDDA 729
QY 855 -KDIAFGLRANQTLDELDFSNVLTDAKAKHLQRLQPSCKLQRLQQLVSCGLTSDCCQD 913
Db 730 YEKLCLEFNGYKTIHLLIQGNL-DSMHSLCEVLKPNPACNLKFLSLGSCSTAQAQKWD 788
QY 914 LASVLSASPSLKELDLQNNLDVGVRLLCGLRHHP 949
Db 789 FFPVLKVNQSLIFDLDTNSLLDKSAKLLCNWKEP 824

RESULT 11
Q9EPG7 PRELIMINARY; PRT; 657 AA.
AC Q9EPG7;
DT 01-WAR-2001 (TREMELrel. 16, Created)
DT 01-WAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-WAR-2003 (TREMELrel. 23, Last annotation update)
DE RNI-like protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvJ;
RX MEDLINE=21310002; PubMed=11416212;
RA Lane R.P., Cutforth T., Young J., Athanasiou M., Friedman C.,
RA Rowen L., Evans G., Axel R., Hood L., Trask B.J.;
RT "Genomic analysis of orthologous mouse and human olfactory receptor
RT loci."
RL Proc. Natl. Acad. Sci. U.S.A. 98:7390-7395 (2001).
DR EMBL: AF321233; AAG45188.1; -.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR007111; NACHT_NTPase.
DR PROSITE; PS50503; LRR_RI; 1.
DR PROSITE; PS50837; NACHT; 1.
SQ SEQUENCE 657 AA; 75410 MW; F4BED9E4BA19AAF3 CRC64;

Query Match 9.0%; Score 676; DB 11; Length 657;
Best Local Similarity 27.1%; Pred. No. 4.8e-42;
Matches 194; Conservative 95; Mismatches 220; Indels 206; Gaps 11;

QY 314 IRDLFGPLDTPQPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGDRFOHFVYFSCREL 372
Db 78 LQHIHDEDIQISEADPTVVLQGAAGIGKTLTKKAVLWADGNLY-QQTHVYFLNGKEI 136
QY 373 AQSKVSVLAEIKGDTATPAPIRQILSRPERLLFILDGVDEPGVWLQEPSSSELCHWSQ 432
Db 137 SQVKEKSPAQLISKHWPSPSEGLEQ-----GQVEKGRVEE----- 172
QY 433 POPADALLGSLGKTLTPEASFLITARTALONLIPSEARVVEVIFGSESSKEYFYR 492
Db 173 ----- 172
QY 493 YFTDERQAIRAFRLVKSNNKELWALCLVPWVSWLACTCLMQMKRKEKLTLSKTTTLCL 552
Db 173 -----TKTSTALFT 182
QY 553 HYLAQALQAQPLG-----PQLRDLCSLAEGIWQKTLFSPDDLRKHGLDGAISTFL 605

```

```
Db 183 YYICSLFPRIPVGCVTLPNETLLRSCKAAVEGIWTKHVLYQONLRKHELTRDILLFL 242
Qy 606 KMGILQEH-PIPLSYSFHLCQFQFFAAMSIVL-EDEKGRGKHSNCIIDLKTLAYAGTH 663
Db 243 DAKVLOQDTEYENCYMFTHLVQFQFFAALFYLLRENLEQDYPSEPFENLYLLESNHTH 302
Qy 664 GLFGASTTRFLGLLSDDEGEREMENIFHCRLSQ--GRNLMOMVPSLQ---LLOPHSLLE 717
Db 303 DPHLEQMKCFGLLNKDRVRQLEETFNLTISNEVREELIACLEGLEKDDSSLSQRFD 362
Qy 718 SLHCLYETRNKTLTQVMAHEFMGCMVETMBELLVCTFCIKFSPRHVKQLQIEGRQHS 777
Db 363 LLHCIVETQDQBEFITQALVYFKIIVRVDEEPQLRIYSCLKCHTLKTMRLTARADLN 422
Qy 778 TWSPTMVVLFWRVPVTDATWQILFSLVKVTRNLKELD----- 815
Db 423 MLDTAEMCLEGAQVVIHYWQDLFSVHTNESLIEMDLYESLMDLSELMKILNEELSHPKC 482
Qy 816 -----SGNSLSH-----SAVSLCKTLRRPRCLLETL 842
Db 483 KLOKLIFRSVDFLNGQDFTFLASNKVTHLDLKETDVGWGLKTLCEALKCKGCKLRVL 542
Qy 843 RLACGLTRAEDCKDLAFGRANQTLTDLSPNVLTDAGAKHLQRLRQPSCKLQRLQIV 902
Db 543 RLASCDLNVARQCKLSNALQTRNSLVFLNLSNLSNDGVKSLCEVLENPNSSLERLALM 602
Qy 903 SCGLTSDCCQDLASVLASPSLKELDLOQNLDVGVRLCEGLRHPACKLRIG 957
Db 603 SCVLTSLKACGDLASVLVNNLSWLSLDGHNILDDAGLNLCDALRPNPCHVQRLG 657

RESULT 12
ID Q8CCN1 PRELIMINARY; PRT; 673 AA.
AC Q8CCN1;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE PAAD and NACHT containing protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=olfactory brain;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK032446; BAC27872.1; -.
SQ SEQUENCE 673 AA; 76367 MW; EE773C592BEC7054 CRC64;

Query Match 7.6%; Score 573; DB 11; Length 673;
Best Local Similarity 33.9%; Pred. No. 2.9e-34;
Matches 152; Conservative 81; Mismatches 154; Indels 62; Gaps 14;

Qy 312 IEIRDLFGPGLDT-QBPRVILQGAAGIGKSTLARQVKEAWGRGQLYGDRFOHVFVFSR 370
Db 146 VDVELTFAPAEASYSTPIVWQSGAGTKTTLVKLVQDWSKGKLYPGQDYVVFVSCR 205
Qy 371 ELAQSKVSLAEI-----KGQATPATPIRQILSRPERLLFLDGVDEPGWVLPQSSSL 426
Db 206 EVVLLPKDLENLICWCCGDD---QAPVTEILRQGRLLFLDGYDE---LQKSSR-- 255
Qy 427 CLHWSQPADALLGSLGKTLIPASFLITARTALQNLIPSLQARVAVLGFSESSR 486
Db 256 -----AECVHTLMRREVP-CSLITITRPPALQSLPEMLGERRHVLVLGFSSEER 305
Qy 487 KEYFYRYFTDERQAIRAFRLVSKNELWALCLVPVWSWLACTCLMQMKRKEKLTLSKT 546
```

```
Db 306 EYFSCFTDEQLKNALEFVQNNVLYKACQVPGICWVCSWLKKQVARGQEVSETPSN 365
Qy 547 TTTLCHLYLAQALQAPLGPQ-----LRDLCSLAAREGIWQKKTLSPPDLRKHGLDG 598
Db 366 STDITAVYSTFLPTDGNDSSELTRHKVLSLCSLAAREGMRHQRLLFEEVLRKHGLDG 425
Qy 599 ATISTFLK-----MGILQEHPIPLSYSFHLCQFQFFAAMSIVLED---EKGRGKHSN 648
Db 426 'PSLTAFLANCIDYRAGLGK-----FYSFRHISQFEEFYAMSLVKEDSQSQGQATHKE 479
Qy 649 C--IIDLEKTLAYGIHGLFGASTTRFLGLLSDDEGEREMENIFHCRL--SGRNLMOVM 704
Db 480 VAKLVDPNHEEV-----TSLQFLFDMLKTEGTLISGLKFCFRIAPSVRQDLKHFK 531
Qy 705 PSLQQLLQPHSLESLSHCLYETRNKTLTQ 733
Db 532 EQIEAIKYKRSWDLFSLYDSKIKK-LTQ 559

RESULT 13
ID Q8IWFS PRELIMINARY; PRT; 953 AA.
AC Q8IWFS;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Caspase recruitment domain family, member 4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Strausberg R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC040339; AAH40339.1; -.
SQ SEQUENCE 953 AA; 107671 MW; 0A9DF167BE87E21A CRC64;

Query Match 5.2%; Score 393.5; DB 4; Length 953;
Best Local Similarity 24.0%; Pred. No. 1.8e-20;
Matches 214; Conservative 131; Mismatches 336; Indels 209; Gaps 38;

Qy 237 VVGTPQAHTSLQPHHPWEPVSRESLQSTWPKNEDFNQKFTQLLLQRPSPQDPLV 296
Db 120 VVNTDPVSRYTQQLRHHLGRDS-KEVLC-----YAKK--BELLLEIYMDTIMELV 167
Qy 297 KRSWPDYVVEENRGLIETRDIFG--PGLDTPRIVILQGAAGIGKSTLARQVKEAWGRG 354
Db 168 -----GFSNESLSGLNSLACLLDHTTGLNQGGETIFILGDAGVGKSMQLQRLQSLWATG 222
Qy 355 QL-YGDRFOHVFYFSCRELA---QSKVVSALAEIKD---GTATPATPIRQILSR-PERLL 406
Db 223 RLDAGVKF--FFHFRCRMFSCKESDRLCQDLQKLYCYPERDPEEVFAFLRPPHVAL 280
Qy 407 FLDGVDEPGWVLPQSSSELCLH-----WSQPADALLGSLGKTLIPASFLITA 458
Db 281 FTFDGLD-----ELHSDLDLSRVPDSSCPWEPNPL-VLLANLSLKLKAGAKLITA 332
Qy 459 RTTALQNLIPSLQARVAVLGFSESSSRKEYFYRYFTDERQAIRAFRLVSKNELWALCL 518
Db 333 RTGI---EVPRQFLRKKVLLRGSFSPSLRVARRMFPERALQDRLLSLEAPNLCSLCS 389
Qy 519 VPVWSWLACTCLMQ-----OMKRKEKLTLSKTTTTL 550
Db 390 VPLFCWILFRCFQHPRAAFEGSPQLPDCDTMTLTDVFLLVTEVHLNRMPQSSIVQRT--- 446
Qy 551 CLHLYLAQALQAPLGPQLRD-LCSL---AAGIQKKTLSPPDLRKHGLDGAISTFLK 606
Db 447 --HSPVETLHAG-----RDTLCSLGQVAHRGMEKSLVFQEEVQASGLQ-----ERDMQ 494
Qy 607 MGILQEHF-----IPLSYSFHLCQFQFFAAMSIVLEDEKGR-----GK 645
```

```

Db 495 LGFLRALPELPGGDOQSQEFFHLLTLCQAEFTAFPLVLDVRVGTQELLRRFFQEMMPAGAA 554
QY 646 HSNCI-----IDLEKTLAYGTHGLFGASTTRFLGLLSDGE-----683
Db 555 TTSYCYPFPLPQCLQSGPAREDLFNKD-----HFQF-----TNLFLCGLLSKAKQKLLRHL 607
QY 684 -----REMENIFHCRLSQGRNLMQWPSL-----QLLOPHSLSLHCLYETRNTKF 730
Db 608 VPAALRRKRKALWAHLFSRLGYSKSLPRVQVESFNQVQAMPTFTWMLRCIYETQ-----663
QY 731 LTQVMAHFPEMGVETDMLLVCTFC-----IKFSRH--VKKLOI-----IEGR 773
Db 664 -SQVQGLAARGICANY-LKLTTCNACSDACSALSFLVHFPKRLALDLDNNLNLDYGV 721
QY 774 QHRSTWSPWVVLFRVVPVTDAYWOILFVSVLKVTRNLKELDLSGNSLSHSAVSKLCKTLR 833
Db 722 ELQPCFSRLTVLRSNVQITDGVKVLSEBELTKYKIVTVLGLYNNQITDVGARYVTKILD 781
QY 834 RPRCLLETLRACGLTAEDCKDLAFGLRANOTLTLDLSFNVLTDAGAKHLQRLR-OP 892
Db 782 ECKGLTH-LKLGKNTITSEGGKYLALAVKNKSISEVGMWGNQVGDGAKAFAEALRNHP 840
QY 893 SKLQRLQLVSCGLTSDCCODLASVLSASPSLKELDLQNNLDDVGRVLLCEGLR-HPAC 951
Db 841 S--LTTLSLANSIGISTEGGKSLARALQONTLSLEILWLTQNELNDEVAESLAEMLVNQT 898
QY 952 KLIRGLDOTT-----LSDMR-----QELRALEQEK 978
Db 899 KHLWLIONQITAKGTQALADALQNSGTGITECLNGNLIRPEEAKVYEDEK 948

RESULT 14
Q8K3Z0 PRELIMINARY; PRT; 1020 AA.
AC Q8K3Z0;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CARD15.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA Iwanaga Y., Davey M.P., Martin T.M., Planck S.R., DePriest M.L.,
RA Baugh M.M., Suing C., Rosenbaum J.T.;
RT "Cloning, sequencing and expression analysis of the murine Nod2/Card15
RT gene.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF520774; AA076073.1;
DR MGD; MGI:2429337; Card15.
DR InterPro; IPR001315; CARD.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR007111; NACTH_WTPase.
DR InterPro; IPR001130; TATD_DNase.
DR PROSITE; PS50209; CARD; 2.
DR PROSITE; PS50503; LRR_R1; 1.
DR PROSITE; PS50837; NACTH; 1.
DR PROSITE; PS01137; TATD; 1; 1.
DR PROSITE; PS01137; TATD; 1; 1.
DR SEQUENCE 1020 AA; 113561 MW; 25504905ECF70FBB CRC64;

Query Match
Best Local Similarity 25.2%; Pred. No. 2.5e-20;
Matches 186; Conservative 122; Mismatches 316; Indels 114; Gaps 28;

QY 311 LIBIRDLFGP-GLDQEPRIVLQAGAGIKSTLQKVKQKAGRGQGLYGRFQHVYFSC 369
Db 255 ILGLELDLFDTHGHLNRDADTILVVGAGSGKSTLQRLHLLWATGRSF-QEFLFIPIFSC 313
370 REL-AQSKVVSALAEIGK-----DGTATPAPIRQIILSRPERLLFILDGVDPGWLQEPSS 424
Db 314 ROLQCVAKPLSLRTLLFHFHCCWPDVAQDDVFOFLDHPDRVLLTFDGLDEFFKFRFTDRR 373
QY 425 ELCLHWS--QPQADALLGSLGKLTLPASFLITARTTALQNLIPSLQARWVEVLGFS 482
Db 374 ----HCSPIDFTSVQTLFNLQGNLLKNACKVLTSPDVAVALURKFFVTE-LQJUGFS 428
QY 483 ESSRKEYFYRYFTDRQAIRAFRLVKSKELWALCLVPWVSWLACTCLMQOMKRREKLT 542
Db 429 EEGIQLYLRKHREPCGVADRLIQLICATSAHLGHLCHLPVFSMMVSRCHRELLLNQNGPPT 488
QY 543 TSKTTTTLCL-HYLAALQAOQP-----LGP-----QLRDLCSLAAGIWQKTLF 586
Db 489 TSDMYLLITLQHLF---LHASPDSPGLGLGSLQSRSLTLLHGLHALRGLAMSCYVF 545
QY 587 SPDDLKRGHGLDCAITSTFLKMGIL--QEHPIPLS---YSFIHLCPQEFFFAANSYVLEDEK 641
Db 546 SAQQLQAAQVSDDIS---LGLFVRAQSSVPGSKAPLEFLHITFOCFFAAFYLAVSADT 601
QY 642 -----GRGKHSNCI--IDLEKTLAYGI-----HGLFGASTTRFLL 675
Db 602 SVASLKHLPSCGRLGSSLLGRLLPNLICQSRVKKGSEAAALQKAEPHNL--QITAAFLA 659
QY 676 GLLSDEGEREMENI FHCRLSQGRNLMQWVPSLQLLQPHSL-----ESLH 720
Db 660 GLLS---QQRDLAALACQVSE-RVLLQROARARSCL-AHSLREHFHSIPPAVPGETKSMH 714
QY 721 C-----LYETRNKTFLTQVMAHFE-----EMGMCVETDMLLVCTFCIKFSRHVK 766
Db 715 AMPGFIWLIRSLYEMOEQLAQEAVERRLDIGHLKLTCFVGPAAECALAFVLOHLORPVA 774
QY 767 LOL-----IEGRHRSWTSPWVLPFRWPVTDAYWOILFVSVLKVTRNLKELDLSGNS 819
Db 775 LQLDVNSGDVGVQVQLRPGCVCTALYLRDNNISDRGARTLVECALRCEQLQKALFNKX 834
QY 820 LSHSAVSKSLKTLRRPRCLLETLRACGLTAEDCKDLAFGLRANOTLTLDLSFNVLTD 879
Db 835 LTDACACSMAKLLAHKONFL-SLRVGNNHITAAAGAEVLAQGLKSNSTSLKFLGFWGNSVD 893
QY 880 AGAKHLQRLRQPSCKLQRLQLVSCGLTSDCCODLASVLSASPSLKELDLQNNLDDVG 939
Db 894 KGTQALAEVYADHQ-NLKWLSLVGNNGSMGAEALALMLEKNKSLKEELCLEENHICDEGV 952
QY 940 RLLCEGL-RHPACKLIRL 956
Db 953 YSLAELGKRNSTLKFLKL 970

RESULT 15
Q9BY26 PRELIMINARY; PRT; 287 AA.
AC Q9BY26;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Leucine-rich-repeat protein RNO2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Shami P.J., Kanai N., Wang L.Y., Vreeze T.M., Parker C.J.;
RT "Identification and characterization of a novel gene that is
RT upregulated in leukemia cells by nitric oxide.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF231021; AAK14942.1;
DR HSSP; P13489; 1A4Y.
DR InterPro; IPR007091; LRR_RNinh.
DR PROSITE; PS50503; LRR_R1; 2.
DR PROSITE; PS50503; LRR_R1; 2.
DR SEQUENCE 287 AA; 31765 MW; BD3816C3255B2F9E CRC64;

```

Query Match	5.1%;	Score 383;	DB 4;	Length 287;
Best Local Similarity	43.7%;	Pred. No. 1.7e-20;		
Matches	93;	Conservative 24;	Mismatches 94;	Indels 2; Gaps 1;
QY	769	LI EGRQHRSTWSPTMVVLFWVPVTDAYWQILFVILKVTNRNLKELDLSGNSLSHSAVKSL	828	
DB	73	LC EGLRH PQ--CR LQMIQLRKQ CLESGACQEMASVLGTNPHELDLTGNAEDLGLRL	130	
QY	829	CKTILRP RCLLETILRLAGCGLTAECDKDLAFGLRANQTLTLELDSFNVLTDAGAKHL	QCR	888
DB	131	CQGLRH P VCR LRTIWLKICRLTAAACDELA STLSVQSLRELDLSL NELGDLGVLL	CEG	190
QY	889	L R P SCKLQRLQVLS CGLTSCDCODLASVLSASPSKELDLQQNNLDDVGVRLL	CEGLRH	948
DB	191	LRHPTCKLQTLRLDSCGLTAKACENLYFTLIGINQTLTDLYTNNALGDTGVRL	CKRLSH	250
QY	949	PACKLIRGLDQTTLLSDEMQRQLRALEQEKPOL	981	
DB	251	PGCKRLVLFQGMDLNKMTHSRRLAALRVTKPYL	283	

Search completed: January 29, 2004, 13:48:11
Job time : 84.1934 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 13:11:31 ; Search time 28.1576 Seconds
(without alignments)
2147.276 Million cell updates/sec

Title: US-09-996-617-2
Perfect score: 7534
Sequence: 1 MAGGAWRLACYLEFLKKEE.....HLIMELWEKSKKGLPLSS 1429

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Issued Patents AA:*
- 1: /cgn2_6/prodata/1/iaa/5A COMB.pap.*
 - 2: /cgn2_6/prodata/1/iaa/5B COMB.pap.*
 - 3: /cgn2_6/prodata/1/iaa/6A COMB.pap.*
 - 4: /cgn2_6/prodata/1/iaa/6B COMB.pap.*
 - 5: /cgn2_6/prodata/1/iaa/PCUS COMB.pap.*
 - 6: /cgn2_6/prodata/1/iaa/backfiles1.pap.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	388.5	5.2	953	4	US-09-999-041A-8
2	388.5	5.2	953	4	US-09-245-281-8
3	388.5	5.2	953	4	US-09-207-359B-8
4	388.5	5.2	953	4	US-09-340-620A-8
5	363.5	4.8	456	2	US-08-910-731-8
6	361.5	4.8	456	2	US-08-910-731-4
7	361.5	4.8	456	2	US-08-795-395-4
8	357	4.7	461	2	US-08-910-731-6
9	354.5	4.7	456	2	US-08-910-731-2
10	354.5	4.7	456	2	US-08-795-395-2
11	348	4.6	953	4	US-09-245-281-43
12	348	4.6	953	4	US-09-207-359B-43
13	348	4.6	953	4	US-09-340-620A-43
14	345	4.6	1130	2	US-08-519-547A-6
15	326.5	4.3	966	4	US-09-207-359B-47
16	253.5	3.4	193	4	US-09-340-620A-61
17	250.5	3.3	195	4	US-09-340-620A-49
18	218.5	2.9	490	4	US-09-999-041A-26
19	218.5	2.9	490	4	US-09-245-281-26
20	218.5	2.9	490	4	US-09-207-359B-26
21	218.5	2.9	490	4	US-09-340-620A-26
22	216	2.9	71	4	US-09-340-620A-58
23	209	2.8	70	4	US-09-340-620A-57
24	209	2.8	70	4	US-09-340-620A-66
25	198	2.6	1151	3	US-08-836-134-23
26	198	2.6	1151	4	US-09-493-784-23
27	198	2.6	1232	3	US-08-836-134-2

28	198	2.6	1232	4	US-09-493-784-2	Sequence 2, Appli
29	179.5	2.4	483	4	US-09-904-615-154	Sequence 154, App
30	175.5	2.3	200	4	US-09-099-041A-11	Sequence 11, Appl
31	175.5	2.3	200	4	US-09-245-281-11	Sequence 11, Appl
32	175.5	2.3	200	4	US-09-207-359B-11	Sequence 11, Appl
33	175.5	2.3	200	4	US-09-340-620A-11	Sequence 11, Appl
34	165	2.2	1466	4	US-09-252-991A-30085	Sequence 30085, A
35	147.5	2.0	793	3	US-09-012-710-10	Sequence 10, Appl
36	147.5	2.0	793	4	US-09-556-273-10	Sequence 10, Appl
37	147	2.0	2482	1	US-08-328-254-6	Sequence 6, Appli
38	143.5	1.9	4302	4	US-09-052-469-8	Sequence 8, Appli
39	143.5	1.9	4302	4	US-08-422-582-8	Sequence 8, Appli
40	143.5	1.9	4339	4	US-09-052-469-6	Sequence 6, Appli
41	143.5	1.9	4339	4	US-08-422-582-6	Sequence 6, Appli
42	142.5	1.9	4302	3	US-08-658-136-5	Sequence 5, Appli
43	142.5	1.9	4303	2	US-08-460-751-2	Sequence 2, Appli
44	138.5	1.8	794	1	US-08-393-333-2	Sequence 2, Appli
45	138.5	1.8	794	3	US-09-087-465-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-09-099-041A-8
; Sequence 8, Application US/09099041A
; Patent No. 6340576
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-076001
; CURRENT APPLICATION NUMBER: US/09/099,041A
; CURRENT FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 953
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-099-041A-8

Query Match	5.2%	Score 388.5;	DB 4;	Length 953;
Best Local Similarity	24.6%	Pred: No. 3.7e-28;		
Matches	217;	Conservative 129;	Mismatches 343;	Indels 193;
Gaps	39;			
QY	237	VYGTTPQAHTSLQPHHPWEPVSRESLCSTWPKNEDFNQKFTQLLLQRPSPQDPLV	296	
Db	120	VVNTDPVSRYTQQLRHHLGRDS-KFVLC-----YAQK-EELLLEEIVMDTLMELV	167	
QY	297	KRSWPDYVEENRGLHIEIRDLFG--PGLDTQPRVILQGAAGIGKSTLARQVKEAWGRG	354	
Db	168	-----GFSNESGLNSLACLLDHTTGILNEOGETIFILGDAGVGSMLQLQSLWATG	222	
QY	355	QL-YGDRFQHVYFSCRELA-----QSKVYSLAELIGKD---GTATAPRIQILSR-PERLL	406	
Db	223	RUDAGVKF--FFHRCRMFSCFKESDRICLQQLLFKHYCYPEDPEEVFAFLRFFHVAL	280	
QY	407	FILDGVDEPGVQLQBPSSSLCLH-----WSQPQADALLGSLGKLTILPEASPLITA	458	
Db	281	FTFDGLD-----ELHSDLDLSRVPDSSCPWPAHPL-VLLANLLSGLLKGASKLLTA	332	
QY	459	RTTALQNLIPSEARWVEVLGFSRSSRKEYFYRFTDERQAIRAFRLVKSNNKELWALCL	518	
Db	333	RTGI---EVPRQFLRKKVLLRGFSPLKRAYARRMPPERALQDRLLSLEAPNLCSLCS	389	
QY	519	VPVWSVLACTCLMQQMKKE-----KLTLSKTTTTLCLHY-----LAQALQAO	563	
Db	390	VPLFCWIIIFRCQHFHFAAFEGSPQLPDCDTMTLTDVLLVTEVHLNRMQSSILVQRNTRSP	449	
QY	564	L-----GPQLRD-LCSL-----AAEGIQWQKTLFSPDLRKHGLDGAISTFLKMGILQHP	614	

Db 450 VETLHAG---RDTLCSLQVAHGMKSLFVFTQBEVQASGLQ-----ERDMQLGFLRALP 502
QY 615 -----IPLSYFHLFCFQEFPAAMSYYLEDEKGR-----GKHSNCI----- 650
Db 503 ELPGGDQOQSYEFFHLTLQAFPTAFVLDLDRVGTQELLRFFQEMWMPAGAAATTCYPPF 562
QY 651 -----IDLEKTEAVGIHGLFGASTTFLGLLSDGE-----R 684
Db 563 LPFQCLQSGPAREDLPKND-----HFQF-----TNLFLCGLLSKAKOKLRLHVPAAALRR 615
QY 685 EMENIFHRLSOGRLNMQWVPSL-----QLLQPHSLSLHCLYETRNKTLFLOVMAHF 738
Db 616 KRKALWAHLFSSLRGLVKSIPRVQVESFNQVQAMPTFIWMLRCIYETQ-----SOKVGQL 670
QY 739 EEMGMCVETDMELLVCTFC-----IKFSRH---VKKLQ-----IEGRQHRSTWSP 781
Db 671 AARGICANY-LKLTVCNACSADCSALSFVHLHHPKRLALDLDNNLNDYGVRELQPCFSR 729
QY 782 TMVLFRRVPTDAYWQILFVSLKVTNKLKELDLGNSLSHSAVKSCLKTLRPRCLLET 841
Db 730 LTVLRUSVNOITDGGVKVISELTKYKIVTYLGLYNNQITDVGARYVTTKILDECKGLTH- 788
QY 842 LRLAGCGLTAEDCKDLAFGLRANQTLTDLDFNVLTDAKAKHLQORLR-QPSCKLQRLQ 900
Db 789 LKLGKNTITSEGGKYALAVKNSKSISEVGMWGNQVDEGAKAFAEALRNHPS--LTTLS 846
QY 901 LVSCGLTSDCCQDLASVLSASPSLKELDLQNNLDDVGVRLICEGLR-HPACKLIRLGLD 959
Db 847 LASNGISTEGGKSLARALQOQNTSLEILWLTQNELNDEVAESLAEMLKVNQTLKHLWLION 906
QY 960 QTT-----LSDEMR-----QELRALEQEK 978
Db 907 QITAKGTAQALADALQOQNTGITEICLNGNLKPEEAKVYEDEK 948

RESULT 2

US-09-245-281-8
; Sequence 8, Application US/09245281
; Patent No. 6369196
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY
; FILE REFERENCE: 07334/118001
; CURRENT APPLICATION NUMBER: US/09/245,281
; EARLIER FILING DATE: 1999-02-05
; EARLIER FILING DATE: 1998-12-08
; EARLIER FILING DATE: 1998-06-17
; EARLIER FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 953
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-245-281-8

Query Match 5.2%; Score 388.5; DB 4; Length 953;
Best Local Similarity 24.8%; Pred. No. 3,7e-28;
Matches 217; Conservative 129; Mismatches 343; Indels 193; Gaps 39;
QY 237 VVGTTPQAHTSLQPHHHPWEPVSRESLCSSTWPKNEDFNQKFTQLILLQORPHPRSDPIV 296
Db 120 VVNTDPVSRYTQOLRHHLGRDS-KFVLC-----YAKQ--EELLLEEYMDIMELV 167
QY 297 KRSWPDYVEENRGLHIEIRDLFG--PGLDTQEPRIIVLOGAIGKSTLQARQVKAHNG 354
Db 168 -----GFSNESLSLSLACLDDHTTGIILNEQOETIFILGDAGVGKSMLLQRLQSLWATG 222

QY 355 QL-YGDRFQHFVYFSCRELA---OSKVVSLAELGKD---GTATPAPIRQILSR-PERLL 406
Db 223 RLDAGVKF--FFHRCRMFSCFKESDRCLQDLLFKHYCYPERDPEEVFAFILLRPPHVAL 280
QY 407 FILDGVDPGWVLPQPSSELCLH-----WSQOPADALLGSLLGKTLIIPASFLITA 458
Db 281 FTFDGLD-----ELHSDLDLSRVPDSSCPWEPAHPL-VLLANLLSGKLLKGASKLITA 332
QY 459 RTTALQNLIPSLQARWVEVLGFSSESSRKEYFYFTDERQAIRAFRLVKNKELWALCL 518
Db 333 RTGI---EVPQFURKKVLLRGFSFHLRARRMFFPERALQDRLLSQLEANPNLCSLCS 389
QY 519 VPVSWLACTCLMQOMKKE-----KLTLTSKTTTTLCLHY-----LAQALQAP 563
Db 390 VPLFCWIIIFRCFQHFRAAFEGSPQLPDCMTWLTDFVLVTEVHLNRMQPSLSIVQNRTP 449
QY 564 L-----GPQLRD-ICSL---AAGIWKQKTLFSPDDLRKGLDGAIISTFLKMGILQHP 614
Db 450 VETLHAG---RDTLCSLQVAHGMKSLFVFTQBEVQASGLQ-----ERDMQLGFLRALP 502
QY 615 -----IPLSYFHLFCFQEFPAAMSYYLEDEKGR-----GKHSNCI----- 650
Db 503 ELPGGDQOQSYEFFHLTLQAFPTAFVLDLDRVGTQELLRFFQEMWMPAGAAATTCYPPF 562
QY 651 -----IDLEKTEAVGIHGLFGASTTFLGLLSDGE-----R 684
Db 563 LPFQCLQSGPAREDLPKND-----HFQF-----TNLFLCGLLSKAKOKLRLHVPAAALRR 615
QY 685 EMENIFHRLSOGRLNMQWVPSL-----QLLQPHSLSLHCLYETRNKTLFLOVMAHF 738
Db 616 KRKALWAHLFSSLRGLVKSIPRVQVESFNQVQAMPTFIWMLRCIYETQ-----SOKVGQL 670
QY 739 EEMGMCVETDMELLVCTFC-----IKFSRH---VKKLQ-----IEGRQHRSTWSP 781
Db 671 AARGICANY-LKLTVCNACSADCSALSFVHLHHPKRLALDLDNNLNDYGVRELQPCFSR 729
QY 782 TMVLFRRVPTDAYWQILFVSLKVTNKLKELDLGNSLSHSAVKSCLKTLRPRCLLET 841
Db 730 LTVLRUSVNOITDGGVKVISELTKYKIVTYLGLYNNQITDVGARYVTTKILDECKGLTH- 788
QY 842 LRLAGCGLTAEDCKDLAFGLRANQTLTDLDFNVLTDAKAKHLQORLR-QPSCKLQRLQ 900
Db 789 LKLGKNTITSEGGKYALAVKNSKSISEVGMWGNQVDEGAKAFAEALRNHPS--LTTLS 846
QY 901 LVSCGLTSDCCQDLASVLSASPSLKELDLQNNLDDVGVRLICEGLR-HPACKLIRLGLD 959
Db 847 LASNGISTEGGKSLARALQOQNTSLEILWLTQNELNDEVAESLAEMLKVNQTLKHLWLION 906
QY 960 QTT-----LSDEMR-----QELRALEQEK 978
Db 907 QITAKGTAQALADALQOQNTGITEICLNGNLKPEEAKVYEDEK 948

RESULT 3

US-09-207-359B-8
; Sequence 8, Application US/09207359B
; Patent No. 6469140
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; FILE REFERENCE: 07334-112001
; CURRENT APPLICATION NUMBER: US/09/207,359B
; CURRENT FILING DATE: 1998-12-08
; PRIOR FILING DATE: 1998-06-17
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 953
; TYPE: PRT

ORGANISM: Homo sapiens
US-09-207-359B-8

Query Match 5.2%; Score 388.5; DB 4; Length 953;
Best Local Similarity 24.6%; Pred. No. 3.7e-28;
Matches 217; Conservative 129; Mismatches 343; Indels 193; Gaps 39;
QY 237 VVGTPQAHTSLQPHHPWPSVRESLCTWPKNEDFNQKFTQLLLQRPHPRSQDPLV 296
DB 120 VVNTDPVSRYTQQLRHHLGRDS-KFVLC-----YAQK-EELLLEEIYMDTIMELV 167
QY 297 KRSWPDYVENRCHLIEIRDLEG--PCLDTQPRVILQGAAGICKSTLARQVKEAWGRG 354
DB 168 -----GFSNESLGSUNSLACLDDHTTGILNEQGETIFILGDAGVGKSMQLQLSLWATG 222
QY 355 QL-YGDRFQHFVYFSCRELA---QSKVWSLAELIGKD---GTATPAPIRQILSR-PERLL 406
DB 223 RLDAGVKF--FPHFRMRMFCSEKESDRCLQDLDFKHICYPERDEEVPFALRPFHVAL 280
QY 407 FILDGVDEPGWVLOQFSSBELCLH-----WSQPQADALLGSLGKLTILPEASFLITA 458
DB 281 FTFDGLD-----ELHSDLDLSRVPDSSCPWEPAPHL-VLLANLLSGKLLGASKLLTA 332
QY 459 RTALONLIPSLQARWVEVLGFSSESRKEYFYRYFTDERQAIRAFRLVKSNNKELWALCL 518
DB 333 RTGI---EVPROFLRKKVLLRGFSPSHLRAYARRMPPERALQDRLLSQLEANPNLCSLCS 389
QY 519 VPVWSWLACTCLMQQMKRKE-----KLTLTSTKTTTLCCLHY-----LAQALQAOP 563
DB 390 VPLFCWIIIPRCQHFRAAFEGSPQLPDCWTMTLTDVFLVTEVHLNMQSSLVQRNTRSP 449
QY 564 L-----GPOLRD-LCSL---AAEGIWQKKTFLSPDDLRKHGLDGAIISTFLKMGILQEH 614
DB 450 VETLHAG---RDTLCSLGQVAHGRMEKSLFVFTQBEVQASGLQ-----ERDMQLGFLRALP 502
QY 615 -----IPLSYFIHLCFQEFPAAMSIVLEDEKGR-----GKHSNCI----- 650
DB 503 ELPGGDQGSYEFPHLTQAFFTAFFVLDDRVGTQELLRFQEWMPAGAAATTSYPPF 562
QY 651 -----IDLEKTLVAYGIHGLFGASTTRFLLGLLSDEGE-----R 684
DB 563 LPFQCLQSGGPAREDLFKNKD---HFQF---TNLFLCGLLSKAKQKLLRHLVPAALRR 615
QY 685 EMENIFHCRLSQGRNLMQWVPSL-----QLLLQPHSLSHSLHCLYETRNKTLFTQVMAHF 738
DB 616 KRKALWAHLFSSLRGYSKSLPRVQVESFNQVQAMPTFIWMLRCIYETQ-----SQKVGQL 670
QY 739 EEMGMCVETDMELLVCTFC-----IKFSRH---VKKLQ-----TEGRQHRSTWSP 781
DB 671 AARGICANY-LKLTVCNACSADCSALSFLVHLHPPKRLALDLDNNLNDYGVRELQPCFSR 729
QY 782 TMVVLFRWPVTDAYWQILFVSLKVTNRNLKELDLSGNSLSHSAVSKLCTLRPRCLLET 841
DB 730 LTVLRSLVNQITDGGVKVLSBELTKYKIVTYLGLYNNQITDVGARYVTYKILDECKGLTH- 788
QY 842 LRLAGCGLTAEDCKDLAFGLRANOTLTDLDSFNVLTDAGAKHLCORLR-QPSCKLQRLQ 900
DB 789 LKLGKNNITSEGGKYLAVALVKNKSISEVGMWGNQVDEGAKAFAPALRNHPS--LTTLS 846
QY 901 LVSCGLTSDCCQDLASVLSASPSELKELDLQNNLDVGVRLLCGLR-HPACKLIRLGLD 959
DB 847 LASNGISTEGKSLARALQNTSLEILWLTONELNDEVAESLAEMLKVNQTLKHLWLN 906
QY 960 QIT-----LSDMR-----QELRALBQEK 978
DB 907 QITAKGTAQLADALQNTGITEICLNGNLIKPEAKVYEDEK 948

RESULT 4
US-09-340-620A-8
; Sequence 8, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:

APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 953
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-340-620A-8

Query Match 5.2%; Score 388.5; DB 4; Length 953;
Best Local Similarity 24.6%; Pred. No. 3.7e-28;
Matches 217; Conservative 129; Mismatches 343; Indels 193; Gaps 39;
QY 237 VVGTPQAHTSLQPHHPWPSVRESLCTWPKNEDFNQKFTQLLLQRPHPRSQDPLV 296
DB 120 VVNTDPVSRYTQQLRHHLGRDS-KFVLC-----YAQK-EELLLEEIYMDTIMELV 167
QY 297 KRSWPDYVENRCHLIEIRDLEG--PCLDTQPRVILQGAAGICKSTLARQVKEAWGRG 354
DB 168 -----GFSNESLGSUNSLACLDDHTTGILNEQGETIFILGDAGVGKSMQLQLSLWATG 222
QY 355 QL-YGDRFQHFVYFSCRELA---QSKVWSLAELIGKD---GTATPAPIRQILSR-PERLL 406
DB 223 RLDAGVKF--FPHFRMRMFCSEKESDRCLQDLDFKHICYPERDEEVPFALRPFHVAL 280
QY 407 FILDGVDEPGWVLOQFSSBELCLH-----WSQPQADALLGSLGKLTILPEASFLITA 458
DB 281 FTFDGLD-----ELHSDLDLSRVPDSSCPWEPAPHL-VLLANLLSGKLLGASKLLTA 332
QY 459 RTALONLIPSLQARWVEVLGFSSESRKEYFYRYFTDERQAIRAFRLVKSNNKELWALCL 518
DB 333 RTGI---EVPROFLRKKVLLRGFSPSHLRAYARRMPPERALQDRLLSQLEANPNLCSLCS 389
QY 519 VPVWSWLACTCLMQQMKRKE-----KLTLTSTKTTTLCCLHY-----LAQALQAOP 563
DB 390 VPLFCWIIIPRCQHFRAAFEGSPQLPDCWTMTLTDVFLVTEVHLNMQSSLVQRNTRSP 449
QY 564 L-----GPOLRD-LCSL---AAEGIWQKKTFLSPDDLRKHGLDGAIISTFLKMGILQEH 614
DB 450 VETLHAG---RDTLCSLGQVAHGRMEKSLFVFTQBEVQASGLQ-----ERDMQLGFLRALP 502
QY 615 -----IPLSYFIHLCFQEFPAAMSIVLEDEKGR-----GKHSNCI----- 650
DB 503 ELPGGDQGSYEFPHLTQAFFTAFFVLDDRVGTQELLRFQEWMPAGAAATTSYPPF 562
QY 651 -----IDLEKTLVAYGIHGLFGASTTRFLLGLLSDEGE-----R 684
DB 563 LPFQCLQSGGPAREDLFKNKD---HFQF---TNLFLCGLLSKAKQKLLRHLVPAALRR 615
QY 685 EMENIFHCRLSQGRNLMQWVPSL-----QLLLQPHSLSHSLHCLYETRNKTLFTQVMAHF 738
DB 616 KRKALWAHLFSSLRGYSKSLPRVQVESFNQVQAMPTFIWMLRCIYETQ-----SQKVGQL 670
QY 739 EEMGMCVETDMELLVCTFC-----IKFSRH---VKKLQ-----TEGRQHRSTWSP 781
DB 671 AARGICANY-LKLTVCNACSADCSALSFLVHLHPPKRLALDLDNNLNDYGVRELQPCFSR 729
QY 782 TMVVLFRWPVTDAYWQILFVSLKVTNRNLKELDLSGNSLSHSAVSKLCTLRPRCLLET 841
DB 730 LTVLRSLVNQITDGGVKVLSBELTKYKIVTYLGLYNNQITDVGARYVTYKILDECKGLTH- 788

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/910,731
FILING DATE: (Herewith)
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/795,395
FILING DATE: 04-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/794,546
FILING DATE: 03-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/024,057
FILING DATE: 16-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: ESMOND, ROBERT W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.3440003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2500
TELEFAX: 202-371-2560

```

; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 461 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-910-731-6

Query Match          4.7%; Score 357; DB 2; Length 461;
Best Local Similarity 37.9%; Pred. No. 1.2e-25;
Matches 86; Conservative 30; Mismatches 91; Indels 20; Gaps 3;

QY 768 QLIEGQHSRSTWPTMVVLFWRVPTDAYWQI-----LFSVLKVTNRNLKELDSGN 818
DB 246 ELCPLGLHFSSRLRTI-----WEGITAKGGCDLCRVLRAKESKELSLAGN 294
QY 819 SLSHSAVSKLCTLRPRCLLETLRAGCGLTAECDKDLAFGLRANQTTLELDSFNVL 878
DB 295 ELGDEGARLLCETLLEPGCQLESWLKSCSFTAACCPHFSSVLAQNRFLLEQISNNRLE 354
QY 879 DAGAKHLQRLRQPSCKLQRLQVSCGLTSDCCQDLASVLSASPSLKELDLQNNLDDVG 938
DB 355 DAGVRELQGLQPGQSVLRVLMADCDVSDSCSSLAATLLANHSRLRELDLSNNCLGDAG 414
QY 939 VRLLCGLRHPACKLRIGLDQTTLSDEMQRLEALEQEKPOLLIIPS 985
DB 415 ILQVESVRQPGCLLEQLVLDIYWSEEMEDRLQALEKPKSRVVIS 461

RESULT 9
US-08-910-731-2
; Sequence 2, Application US/08910731
; Patent No. 5932440
; GENERAL INFORMATION:
; APPLICANT: CHATTERJEE, DEB K.
; APPLICANT: SHANDILYA, HARINI
; TITLE OF INVENTION: Mammalian Ribonuclease Inhibitors and Use Thereof
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,731
; FILING DATE: (Herewith)
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/795,395
; FILING DATE: 04-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/794,546
; FILING DATE: 03-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,057
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.3440003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 456 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-910-731-2

Query Match          4.7%; Score 354.5; DB 2; Length 456;
Best Local Similarity 34.3%; Pred. No. 2e-25;
Matches 108; Conservative 35; Mismatches 101; Indels 71; Gaps 8;

QY 752 LVCTCFIKFSRHVKQLQLEGRQHSRSTWPTMVVLFWRVPTDAYWQILFSLVKVTRNLK 811
DB 127 LICEGLLPQCHLEKLEK-----EYCRLTAAACEPLASVLRATRAK 168
QY 812 ELDLGNSLSHSAVSKLCTLRPRCLLETLRAGCGLTAECDKDLAFGLRANQ-TLTEL 870
DB 169 EUTVSNNDIGEGARVLCGGIADSAQCLETURLENCGLTPANCKDLC-GIVASQASLREL 227
QY 871 DLSFNVLTDAGAKHLQRLRQPSCKLQRLQVSCGLTSDCCQDLASVLSASPSLKELDLQ 930
DB 228 DLGNSGLDAGIAELCPGLSPASRLKTLMLWECDITASGCRDLCRVLAQAKETLSELSLA 287
QY 931 QNNLDDVGRLLCEGLRHPACK-----LIRLGLDQTT 962
DB 288 GNKLDEGARLICESLLOPGCQLESWLKSCSLTAACQHVSLMLTONKHLLEQLSSNK 347
QY 963 LSDENRQEL-RALEQEKPOLLIIFSRKPSVMTPTTEGLDTGEMSNS-TSSLKQRLGSERA 1020
DB 348 LGDSGIQELCOALSQPGTTLRVLC-----LGDCEVTNSGCSL-----A 386

1021 ASHVAQANLKLIDVS 1035
387 SILLANRSRLRELDLS 401

RESULT 10
US-08-795-395-2
; Sequence 2, Application US/08795395
; Patent No. 5965399
; GENERAL INFORMATION:
; APPLICANT: CHATTERJEE, DEB K.
; APPLICANT: SHANDILYA, HARINI
; TITLE OF INVENTION: Cloning and Expression of Rat Liver and
; TITLE OF INVENTION: Porcine Liver Ribonuclease Inhibitor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
; STREET: 1100 NEW YORK AVE., N.W., SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/795,395
; FILING DATE: 04-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/024,057
; FILING DATE: 16-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0942.3440002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

```

[illegible]

; SEQ ID NO 43
; LENGTH: 953
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-207-359B-43

Query Match 4.6%; Score 348; DB 4; Length 953;
Best Local Similarity 21.6%; Pred. No. 3.4e-24;
Matches 195; Conservative 130; Mismatches 361; Indels 218; Gaps 32;

```
QY 234 WAAVGTTPQAHTSLQPHHPWEPVRESLCTWPKWNEDFNQKFTQLLLQRP----- 287
DB 103 WLSEIGFSP-----SQLIRTKTIIVTDPVSRYTQQLRHQLGRDSKFM 144
QY 288 --HPSQDPLVKESWPD-----YVEENRGHLEIRDLF--GPGLDTOEPRIVILQGA 336
DB 145 LCYAKEDLLLEETVMDTLMGLVGFNNENLGLSLGDLCLDHSSTGVNHEGETVFVFGDA 204
QY 337 GIGKSTLARQVKEAWGRGQLYGDRFOHVYFSCRELA---QSKVSLAEELIGKD---GTA 390
DB 205 GVGKSMLLQRLQSLWASGRLTSTA-KFFHFRCRMFSCFKESDMLSLQDLLFKHFCYPEQ 263
QY 391 TPAPTRQILSR-PERLLFILGVDV---PGWVLOEPPSELCLHWSQPOPADALLGSLGK 446
DB 264 DPEEVFSFLRPHTALTFTDGLDELHSDFDLSRPVDS--CCWPEPAHL-VLLANLLSG 320
QY 447 TILPEASFLITARTTALQNLIPSLQARWVEVLGFSSESRKEYFYRYTDERQAIAPRL 506
DB 321 RLKKGAGKLLTARTGV---EVPRLQRKKVLLRGFSPSHLRAYARRMFPERTAQEHLLQ 377
QY 507 VKSNKELWALCLVPVWSWLAECTCLMQ-----QMKRK 537
DB 378 LDANPNLCSLCGVPLFCWIIIFRCFOHFQTVFEGSSQLPDCAVTLTDVFLVTEVHLNRP 437
QY 538 EKLTLSTKTTTTLCHLYLAQAQALPQPLGRDLCSLAAEGIWOKKTLFSPDDLKRGHD 597
DB 438 QPSSLVQRNTRS-----PAETLRAG--WRTLHALGEVAHRGTDKSLFVFGQEVQASKLO 490
QY 598 GAIISTFLKMGILQEHF-----IPLSYSFHILCFQOEFFAAMSYLEDE----- 640
DB 491 ---EGDLQGLFLRALPDVGPQOGQSYEFFHLLTQAFFTAFLVADKSTRELLRFFRE 546
QY 641 ---KRGKHSNCIIDIETKLEAYGHGLFGAS-----TTRFLLGLLSDEG 682
DB 547 WTSPEGATSSSCHSP-----FSQCLGRSRLGPDPRNKHDPQFTNLFVCGLLAKAR 600
QY 683 EREMENIFCHRLSQGRNLMQW-----VPSL-----QLLQPHSLESILHCLY 723
DB 601 QKLLRLQVPAKILRRKRKALWAHLFASLSRYLSKSLPRVQSGGFNQVHAMPTFLWMLRCIY 660
QY 724 ETRNKTFLTQVMAHPEEMGCVETDMLLAVCTFCIKFSRHVKKIQLI EGRQHRSTWSP 783
DB 661 ETQ-----SOKVGRLAARGI-----SADYLUKLAFACNACSDCSALSFLVHHFRQL----- 706
QY 784 VVLFRRVVPVTDAYWQILFVLKTRNLKELDLSGNSLSHSAVKSCLCKTLRRPRCLLETUR 843
DB 707 -----ALDLDNNLNNDYGVQELQPCFSR-----LTVIR 734
QY 844 LAGCGLTAECDKDLAPGLRANOTLTBLDSFNVLTAGAKHLCQRLQPSCK-LQRLQIV 902
DB 735 LSVNQITDGTGVKVLCELTKYIVTFLGYNQITDIGARYVAQILDE--CRGLKHLKLG 792
QY 903 SCGLTSDCCODLASVLSASPSLKELDLQONNLDVGVRLLCGLR-HPACKLIRLGLDQT 961
DB 793 KNRTISEGCKVALAVNNTSIVDVGWGNQIGDEGAKAFALADKHPS--ITTLSLAPN 950
QY 962 TILSDENRQEL-RALEQEKPLLIIFSRKPSVMTPTTEGLDTGEMSNSTSSLRKQRLGSE 1020
DB 851 GISPEGGKSLAQALKNTTLTIW-----LTKNELNDESACFAEMLRVNQT 897
QY 1021 ASHV 1024
DB 898 LRHL 901
```

RESULT 13

US-09-340-620A-43
; Sequence 43; Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US/09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US/09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US/09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US/09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 953
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-340-620A-43

Query Match 4.6%; Score 348; DB 4; Length 953;

Best Local Similarity 21.6%; Pred. No. 3.4e-24;

Matches 195; Conservative 130; Mismatches 361; Indels 218; Gaps 32;

```
QY 234 WAAVGTTPQAHTSLQPHHPWEPVRESLCTWPKWNEDFNQKFTQLLLQRP----- 287
DB 103 WLSEIGFSP-----SQLIRTKTIIVTDPVSRYTQQLRHQLGRDSKFM 144
QY 288 --HPSQDPLVKESWPD-----YVEENRGHLEIRDLF--GPGLDTOEPRIVILQGA 336
DB 145 LCYAKEDLLLEETVMDTLMGLVGFNNENLGLSLGDLCLDHSSTGVNHEGETVFVFGDA 204
QY 337 GIGKSTLARQVKEAWGRGQLYGDRFOHVYFSCRELA---QSKVSLAEELIGKD---GTA 390
DB 205 GVGKSMLLQRLQSLWASGRLTSTA-KFFHFRCRMFSCFKESDMLSLQDLLFKHFCYPEQ 263
QY 391 TPAPTRQILSR-PERLLFILGVDV---PGWVLOEPPSELCLHWSQPOPADALLGSLGK 446
DB 264 DPEEVFSFLRPHTALTFTDGLDELHSDFDLSRPVDS--CCWPEPAHL-VLLANLLSG 320
QY 447 TILPEASFLITARTTALQNLIPSLQARWVEVLGFSSESRKEYFYRYTDERQAIAPRL 506
DB 321 RLKKGAGKLLTARTGV---EVPRLQRKKVLLRGFSPSHLRAYARRMFPERTAQEHLLQ 377
QY 507 VKSNKELWALCLVPVWSWLAECTCLMQ-----QMKRK 537
DB 378 LDANPNLCSLCGVPLFCWIIIFRCFOHFQTVFEGSSQLPDCAVTLTDVFLVTEVHLNRP 437
QY 538 EKLTLSTKTTTTLCHLYLAQAQALPQPLGRDLCSLAAEGIWOKKTLFSPDDLKRGHD 597
DB 438 QPSSLVQRNTRS-----PAETLRAG--WRTLHALGEVAHRGTDKSLFVFGQEVQASKLO 490
QY 598 GAIISTFLKMGILQEHF-----IPLSYSFHILCFQOEFFAAMSYLEDE----- 640
DB 491 ---EGDLQGLFLRALPDVGPQOGQSYEFFHLLTQAFFTAFLVADKSTRELLRFFRE 546
QY 641 ---KRGKHSNCIIDIETKLEAYGHGLFGAS-----TTRFLLGLLSDEG 682
DB 547 WTSPEGATSSSCHSP-----FSQCLGRSRLGPDPRNKHDPQFTNLFVCGLLAKAR 600
QY 683 EREMENIFCHRLSQGRNLMQW-----VPSL-----QLLQPHSLESILHCLY 723
DB 601 QKLLRLQVPAKILRRKRKALWAHLFASLSRYLSKSLPRVQSGGFNQVHAMPTFLWMLRCIY 660
QY 724 ETRNKTFLTQVMAHPEEMGCVETDMLLAVCTFCIKFSRHVKKQLI EGRQHRSTWSP 783
```


QY 926 ELDLQNNLDDVGVRLICEGLR 947
 Db 1076 VMDVQYNKFTAGAAQQAASLR 1097

RESULT 15
 US-09-207-359B-47
 ; Sequence 47, Application US/09207359B
 ; Patent No. 6469140
 ; GENERAL INFORMATION:
 ; APPLICANT: Bertin, John
 ; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
 ; FILE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
 ; FILE REFERENCE: 07334-112001
 ; CURRENT APPLICATION NUMBER: US/09/207,359B
 ; CURRENT FILING DATE: 1998-12-08
 ; PRIOR APPLICATION NUMBER: US 09/099,041
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: US 09/019,942
 ; PRIOR FILING DATE: 1998-02-06
 ; NUMBER OF SEQ ID NOS: 47
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 47
 ; LENGTH: 966
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: majority sequence
 US-09-207-359B-47

Query Match 4.3%; Score 326.5; DB 4; Length 966;
 Best Local Similarity 21.9%; Pred. No. 4.5e-22;
 Matches 187; Conservative 113; Mismatches 325; Indels 227; Gaps 31;

QY 271 NEDFNQKFTQLLLQRP-----HPSQDPLVKSQWPD-----YVEENRGHLEIR 315
 Db 122 NTPDVSRYTQQLRHQLGRDSKFVLCYAKQEDLLEIYMDTLMGLVGFNSLGLGLA 181

QY 316 DLF--CPGLDTPRVLGGAGICKSTLQVKEAWGQGL--YQDRQHVYFSCREL 372
 Db 182 CLLDHTGVNLQGETVFLVGDAGVGSMLLQRLQSLWASGRUTAGAKF--FFHFRCPWF 239

QY 373 A---QSKVWSLAELIGKD---GTATPAPIRQILSR--PERLLFILDGVDFPGWVLOEPSSE 425
 Db 240 SCFESDRLSLQDLLEFKHFCYEPQDPEEVFAFLRPFHVALFTFDGLD-----ELHSD 292

QY 426 LCLH-----WSQPPADALGSLGKTLILPEASFLITARTALQNLIPSLQARWYE 477
 Db 293 LQLSRVPDSSCPWEPAHPL-VLLANLLSGKLLKGAGKLLTARTGV---EVPRQLLRKKVL 348

QY 478 VLGFSSESRKEYFYRYFTDERQAIRAFRLVKSNNELWALCLVPWVSWLACTCLMQ----- 532
 Db 349 LRGFSPSHLRAYARRMFPERRAAQDHLLSLQDANPNLCSCGPVLCFVCIIFRCFQHFQAAF 408

QY 533 -----OMKRKEKLTLSKTTTLTLC--LHYLAQALQAOPLGP 566
 Db 409 EGSSSQLPDCAVTLTDVFLVTEVHLNRMQPSSLVQPNTRSPAETLHAGRDTLHA----- 463

QY 567 QLRLDCLSLAAEGIWQKTLFSPDDLRKGLDGAIIITFLKMGILQHP-----IPLSYS 620
 Db 464 ----LGEVAHRTGDKSLFVFGQEVQASGLQ-----EGDLQLGFRLALPDVPGGQGSYE 515

QY 621 FHLCPQEFPAANSYVLEDEKGRCK-----HSCNIDLEKTLIAYG 661
 Db 516 FPHLTLOAFPTAFPLVADDKVGTFQELLRPFQEWTSFGGAASSCHSSFL-----SPQCLG 570

QY 662 IHGLFCASITRFLGLLSDEGEREMENI FHCRL---SOGNLMQWVPSL----- 707
 Db 571 GSGRAGED-----LFPKNDHPQFTNLVFCGLLAKAKOKLLRQLVPAALRRKRKALWA 623

QY 708 -----QLLQPHSLSLHCLYETRNKTFLTQYMAHFEMGMCV 745

Search completed: January 29, 2004, 13:48:52
 Job time : 31.1576 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 13:46:01 ; Search time 249.019 Seconds
(without alignments)
1192.602 Million cell updates/sec

Title: US-09-996-617-2

Perfect score: 7534

Sequence: 1 MAGGAWGRLACYLEFLKKEE.....HLIMELWEKSGKGLPLSS 1429

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7534	100.0	1429	10	US-09-996-617-2 Sequence 2, Appli
2	7534	100.0	1429	10	US-09-931-071-2 Sequence 2, Appli
3	7534	100.0	1429	12	US-10-028-374-15 Sequence 15, Appl
4	7534	100.0	1429	12	US-10-183-770-15 Sequence 15, Appl
5	7534	100.0	1429	15	US-10-028-392-11 Sequence 11, Appl
6	7502	99.6	1473	10	US-09-388-221-2 Sequence 2, Appli
7	7488	99.4	1429	12	US-10-028-374-3 Sequence 3, Appli
8	7488	99.4	1429	12	US-10-183-770-3 Sequence 3, Appli
9	7364	97.7	1399	10	US-09-388-221-4 Sequence 4, Appli
10	7332	97.3	1443	10	US-09-388-221-6 Sequence 6, Appli
11	6344	84.2	1454	10	US-09-388-221-10 Sequence 10, Appl
12	6174	81.9	1424	10	US-09-388-221-12 Sequence 12, Appl
13	4072	54.0	764	12	US-10-407-866-92 Sequence 92, Appl
14	2171.5	28.8	442	11	US-09-895-298-139 Sequence 139, App
15	1382.5	18.4	1027	12	US-10-407-866-68 Sequence 68, Appl

16	1380	18.3	1061	15	US-10-066-521-18 Sequence 18, Appl
17	1375	18.3	1035	11	US-09-965-621-24 Sequence 24, Appl
18	1375	18.3	1035	12	US-10-407-866-24 Sequence 24, Appl
19	1364	18.1	1004	12	US-10-108-260A-3161 Sequence 3161, Ap
20	1279	17.0	1034	12	US-10-028-374-18 Sequence 18, Appl
21	1279	17.0	1034	12	US-10-132-967-5 Sequence 5, Appli
22	1279	17.0	1034	12	US-10-183-770-18 Sequence 18, Appl
23	1279	17.0	1034	14	US-10-127-516-5 Sequence 5, Appli
24	1279	17.0	1034	14	US-10-027-629-5 Sequence 22, Appl
25	1154	15.3	896	15	US-10-066-521-22 Sequence 22, Appl
26	1037	13.8	994	11	US-09-965-621-16 Sequence 16, Appl
27	1037	13.8	994	12	US-10-407-866-16 Sequence 16, Appl
28	1037	13.8	994	15	US-10-066-521-24 Sequence 24, Appl
29	1008	13.4	919	12	US-10-094-749-2718 Sequence 2718, Ap
30	985.5	13.1	1162	15	US-10-216-645-2 Sequence 2, Appli
31	973.5	12.9	635	12	US-10-407-866-90 Sequence 90, Appl
32	972.5	12.9	858	9	US-09-848-035-8 Sequence 8, Appli
33	972.5	12.9	858	10	US-09-986-224-8 Sequence 4, Appli
34	969	12.9	1143	15	US-10-216-645-4 Sequence 4, Appli
35	948	12.6	1062	12	US-10-239-663-43 Sequence 43, Appl
36	945.5	12.5	1033	12	US-10-132-967-2 Sequence 2, Appli
37	945.5	12.5	1033	14	US-10-127-516-2 Sequence 2, Appli
38	945.5	12.5	1033	14	US-10-027-629-2 Sequence 2, Appli
39	945.5	12.5	1344	15	US-10-066-521-6 Sequence 6, Appli
40	945.5	12.5	1016	10	US-09-986-224-19 Sequence 19, Appl
41	941.5	12.5	1049	12	US-10-239-663-42 Sequence 42, Appl
42	924.5	12.3	732	12	US-10-407-866-72 Sequence 72, Appl
43	922.5	12.2	952	12	US-10-407-866-70 Sequence 70, Appl
44	913.5	12.1	674	12	US-10-407-866-89 Sequence 89, Appl
45	908.5	12.1	980	9	US-09-848-035-13 Sequence 13, Appl

ALIGNMENTS

RESULT 1

US-09-996-617-2
; Sequence 2, Application US/09996617
; Patent No. US20020128198A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; FILE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-340001
; CURRENT APPLICATION NUMBER: US/09/996, 617
; PRIOR FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: 09/931,071
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/428,252
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1429
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-996-617-2

Query Match 100.0%; Score 7534; DB 10; Length 1429;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGGAWGRLACYLEFLKKEELKEFOLLANKAHSRSSGETPAQPEKTSMEVASVYLAQ 60

DB 1 MAGGAWGRLACYLEFLKKEELKEFOLLANKAHSRSSGETPAQPEKTSMEVASVYLAQ 60

QY 61 YGEQRWDIALHTWQMGLRSLCAQAGAGHSPFPYSPSPHGLSPSQPTSTAVLMPW 120

DB 61 YGEQRWDIALHTWQMGLRSLCAQAGAGHSPFPYSPSPHGLSPSQPTSTAVLMPW 120

QY 121 IHELPAGCTQSGERRVLRQLPDTSGRRMREISASLLYQALPSPDPHSPSPSNAPTST 180

```
Db 121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSPDPHESPQSNAPTST 180
Qy 181 AVLGSGSPPOPSLAPREQEAPGTQWPLDETSGIYYTEIREREREKSGRPPWAAVGT 240
Db 181 AVLGSGSPPOPSLAPREQEAPGTQWPLDETSGIYYTEIREREREKSGRPPWAAVGT 240
Qy 241 PPOAHTSLOPHHHHPWSPVRESLCSWPKWNEFNQKFTQLLLQRPSPRSQDPLVKRSW 300
Db 241 PPOAHTSLOPHHHHPWSPVRESLCSWPKWNEFNQKFTQLLLQRPSPRSQDPLVKRSW 300
Qy 301 PDYVEENRGLHIEIRDLFGGLDTQPRIVILQGAAGIKSTLARQVKEAWGRGQLYGDR 360
Db 301 PDYVEENRGLHIEIRDLFGGLDTQPRIVILQGAAGIKSTLARQVKEAWGRGQLYGDR 360
Qy 361 FQHVYFSCRELAQSKVVSIAELIGKGTATPAPIRQILSRPERLLFILDGVDGCVWLQ 420
Db 361 FQHVYFSCRELAQSKVVSIAELIGKGTATPAPIRQILSRPERLLFILDGVDGCVWLQ 420
Qy 421 EPSSELCHWSQOPADALLGSLGKTTILPEASFLITARTTALQNLIPSLEQARWVEVLG 480
Db 421 EPSSELCHWSQOPADALLGSLGKTTILPEASFLITARTTALQNLIPSLEQARWVEVLG 480
Qy 481 FSESSRKEYFYRYFTDERQAIRAFRLVKNKELWALCLVPWVSWLACTCLMQMKRKEKL 540
Db 481 FSESSRKEYFYRYFTDERQAIRAFRLVKNKELWALCLVPWVSWLACTCLMQMKRKEKL 540
Qy 541 TLTSKTTTTLCHLYLAQAQAPLGPQLRDCLSLAAEGIWOKTLPSPDDLRKHGLDGA 600
Db 541 TLTSKTTTTLCHLYLAQAQAPLGPQLRDCLSLAAEGIWOKTLPSPDDLRKHGLDGA 600
Qy 601 ISTFLKMGILQEHPIPLSVSFHLCQEPFAFMSYVLEDEKGRGHSNCIIDLKTELEY 660
Db 601 ISTFLKMGILQEHPIPLSVSFHLCQEPFAFMSYVLEDEKGRGHSNCIIDLKTELEY 660
Qy 661 GIHGLFGASTTFLGLLSDGGEREMENIFHCRLSQGRNLMQWVPSLQLLLOPHSLES 720
Db 661 GIHGLFGASTTFLGLLSDGGEREMENIFHCRLSQGRNLMQWVPSLQLLLOPHSLES 720
Qy 721 CLYETRNKFTLQVMAHFEMGCVETDMELLVCTCFIKFSRHVKKLQIEGRQHRSTWS 780
Db 721 CLYETRNKFTLQVMAHFEMGCVETDMELLVCTCFIKFSRHVKKLQIEGRQHRSTWS 780
Qy 781 PTMWVLFVRVPVTDAYWQILFVSLKYTRNLKELDLSGNSLSHSAVKSCLKTLRRPRCLLE 840
Db 781 PTMWVLFVRVPVTDAYWQILFVSLKYTRNLKELDLSGNSLSHSAVKSCLKTLRRPRCLLE 840
Qy 841 TLRLACGLTAEDCKDLAFGLRANQTLTBLDLSFNVLTDAGAKHLCQRLQPSCKLQRLQ 900
Db 841 TLRLACGLTAEDCKDLAFGLRANQTLTBLDLSFNVLTDAGAKHLCQRLQPSCKLQRLQ 900
Qy 901 LVSCGLTSCCQDLASVLSASPSLKBELDLQONNLDVGVRLLCBGLRHPACKLIRLGLDQ 960
Db 901 LVSCGLTSCCQDLASVLSASPSLKBELDLQONNLDVGVRLLCBGLRHPACKLIRLGLDQ 960
Qy 961 TTLSDSMRBELRALOEKQOLLIFSRKPSVMTPTGLDTGEMNSTSLKQRLGSE 1020
Db 961 TTLSDSMRBELRALOEKQOLLIFSRKPSVMTPTGLDTGEMNSTSLKQRLGSE 1020
Qy 1021 ASHVAQANKLIDVSKIFPIAETAEBSSEPVVPELLCVPSPASQGLDHTKPLGTDGDDFW 1080
Db 1021 ASHVAQANKLIDVSKIFPIAETAEBSSEPVVPELLCVPSPASQGLDHTKPLGTDGDDFW 1080
Qy 1081 GPTGPVATEVWKEKNLYRVHPFVAGSYRWPNTGLCFVNRVATVVEIFCVWDQFLGEIN 1140
Db 1081 GPTGPVATEVWKEKNLYRVHPFVAGSYRWPNTGLCFVNRVATVVEIFCVWDQFLGEIN 1140
Qy 1141 PQHSMWVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
Db 1141 PQHSMWVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
Qy 1201 ELHHIVLENPSPFLGVLKMIHNAALRFIPVTSVLLYHRVHPVEVTFHLYLIPSDCSIR 1260
```

```
Db 1201 ELHHIVLENPSPFLGVLKMIHNAALRFIPVTSVLLYHRVHPVEVTFHLYLIPSDCSIR 1260
Qy 1261 KELELCYRSPGSDQLFSEFYVGHLSGIRLQVKDKKDETLVWEALVKPGDLMPATTLIPP 1320
Db 1261 KELELCYRSPGSDQLFSEFYVGHLSGIRLQVKDKKDETLVWEALVKPGDLMPATTLIPP 1320
Qy 1321 ARIAVPSPLDAPQLLHFVDQYREQRIARVTSVEVVLDKLGQVLSEQEQYVERVLAENTRPS 1380
Db 1321 ARIAVPSPLDAPQLLHFVDQYREQRIARVTSVEVVLDKLGQVLSEQEQYVERVLAENTRPS 1380
Qy 1381 QMRKLFSLSQSWDRCKDGLYQALKETHPHILMELWEKSGKGLPLSS 1429
Db 1381 QMRKLFSLSQSWDRCKDGLYQALKETHPHILMELWEKSGKGLPLSS 1429
```

RESULT 2

US-09-931-071-2

; Sequence 2, Application US/09931071

; Patent No. US20020128219A1

; GENERAL INFORMATION:

; APPLICANT: Bertin, John

; APPLICANT: Alnemri, Emad S.

; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED

; FILE OF INVENTION: PROTEIN FAMILY AND USES THEREOF

; CURRENT APPLICATION NUMBER: 07334-335001

; CURRENT FILING DATE: 2002-03-18

; PRIOR APPLICATION NUMBER: 09/428,252

; PRIOR FILING DATE: 1999-10-27

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 1429

; TYPE: PRK

; ORGANISM: Homo sapiens

; US-09-931-071-2

Query Match 100.0%; Score 7534; DB 10; Length 1429;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 MAGGAWGRLACYLEFLKKEELKEFOLLANKAHSRSSGGETPAQEKTSGMVEVASYLVAQ 60
Db 1 MAGGAWGRLACYLEFLKKEELKEFOLLANKAHSRSSGGETPAQEKTSGMVEVASYLVAQ 60
Qy 61 YGEQAWDLALHTWEQMGRLSICAAQEGAGHSFPYSPSEPHLGSPSQSTSTAVLMPW 120
Db 61 YGEQAWDLALHTWEQMGRLSICAAQEGAGHSFPYSPSEPHLGSPSQSTSTAVLMPW 120
Qy 121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSPDPHESPQSNAPTST 180
Db 121 IHELPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSPDPHESPQSNAPTST 180
Qy 181 AVLGSGSPPOPSLAPREQEAPGTQWPLDETSGIYYTEIREREREKSGRPPWAAVGT 240
Db 181 AVLGSGSPPOPSLAPREQEAPGTQWPLDETSGIYYTEIREREREKSGRPPWAAVGT 240
Qy 241 PPOAHTSLOPHHHHPWSPVRESLCSWPKWNEFNQKFTQLLLQRPSPRSQDPLVKRSW 300
Db 241 PPOAHTSLOPHHHHPWSPVRESLCSWPKWNEFNQKFTQLLLQRPSPRSQDPLVKRSW 300
Qy 301 PDYVEENRGLHIEIRDLFGGLDTQPRIVILQGAAGIKSTLARQVKEAWGRGQLYGDR 360
Db 301 PDYVEENRGLHIEIRDLFGGLDTQPRIVILQGAAGIKSTLARQVKEAWGRGQLYGDR 360
Qy 361 FQHVYFSCRELAQSKVVSIAELIGKGTATPAPIRQILSRPERLLFILDGVDGCVWLQ 420
Db 361 FQHVYFSCRELAQSKVVSIAELIGKGTATPAPIRQILSRPERLLFILDGVDGCVWLQ 420
Qy 421 EPSSELCHWSQOPADALLGSLGKTTILPEASFLITARTTALQNLIPSLEQARWVEVLG 480
Db 421 EPSSELCHWSQOPADALLGSLGKTTILPEASFLITARTTALQNLIPSLEQARWVEVLG 480
```

481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNGKELWALCLVPWVSWLACTCLMQMKKEKL 540
Db
481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNGKELWALCLVPWVSWLACTCLMQMKKEKL 540
Qy 541 TLTSKTTTTLCHYLAQAALQAQPLGQPLRDLCSLAAEGIWOKTFLFSPDDLRKHGIDGAI 600
Db 541 TLTSKTTTTLCHYLAQAALQAQPLGQPLRDLCSLAAEGIWOKTFLFSPDDLRKHGIDGAI 600
Qy 601 ISTFLKMGILQEHPIPLSYFHLFCQEFFAAMSVYVLEDEKGRGHSNCIIIDLEKTLAY 660
Db 601 ISTFLKMGILQEHPIPLSYFHLFCQEFFAAMSVYVLEDEKGRGHSNCIIIDLEKTLAY 660
Qy 661 GIHGLFGASTTRFLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLOPHSLESLSH 720
Db 661 GIHGLFGASTTRFLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLOPHSLESLSH 720
Qy 721 CLYETRNKFTLQVMAHFEMGCMVETDMELLVCTFCIKFSRHVKKLQIEGRQHRSTWS 780
Db 721 CLYETRNKFTLQVMAHFEMGCMVETDMELLVCTFCIKFSRHVKKLQIEGRQHRSTWS 780
Qy 781 PTMVVLFWRVPVTDAYWQILFSLVKVTRNLKELDLSGNSLSHSAVKSCLKTLRRPRCLLE 840
Db 781 PTMVVLFWRVPVTDAYWQILFSLVKVTRNLKELDLSGNSLSHSAVKSCLKTLRRPRCLLE 840
Qy 841 TLRLAGCGLTAECDKDLAFGLRANQTLTDLDSFNVLTDAGAKHLQRLRQPSCKLQRLQ 900
Db 841 TLRLAGCGLTAECDKDLAFGLRANQTLTDLDSFNVLTDAGAKHLQRLRQPSCKLQRLQ 900
Qy 901 LVSCGLTSDCCODLASVLSASPSLSKELDLQNNLDVGVRLLCGHRHPACKLIRLGLDQ 960
Db 901 LVSCGLTSDCCODLASVLSASPSLSKELDLQNNLDVGVRLLCGHRHPACKLIRLGLDQ 960
Qy 961 TTLSDEMROELRALQEKEPOLLIFFSRKPSVMTPTGLDTGEMNSSTSLKRLQSGSRA 1020
Db 961 TTLSDEMROELRALQEKEPOLLIFFSRKPSVMTPTGLDTGEMNSSTSLKRLQSGSRA 1020
Qy 1021 ASHVAQANLLDVSIFPIAEIABESSPEVVPVELLCVPSPASQGDHLTKPLGTDDDFW 1080
Db 1021 ASHVAQANLLDVSIFPIAEIABESSPEVVPVELLCVPSPASQGDHLTKPLGTDDDFW 1080
Qy 1081 GPTGPVATEVDKELNLYRHVFPVAGSYRPNPTGLCFVNRVAVTVIEFCVWDOFLGEIN 1140
Db 1081 GPTGPVATEVDKELNLYRHVFPVAGSYRPNPTGLCFVNRVAVTVIEFCVWDOFLGEIN 1140
Qy 1141 PQHSMMVAGPLLDIKABPGAVEAVHLPFHVALOGCHVDTSLFQMAHFKEGMLLEKPARV 1200
Db 1141 PQHSMMVAGPLLDIKABPGAVEAVHLPFHVALOGCHVDTSLFQMAHFKEGMLLEKPARV 1200
Qy 1201 ELHHIVLENPSFPLGVLKMIHNAIRFIPVTSVLLYHRVHPPEVTFHLYLIPSDCSIR 1260
Db 1201 ELHHIVLENPSFPLGVLKMIHNAIRFIPVTSVLLYHRVHPPEVTFHLYLIPSDCSIR 1260
Qy 1261 KELELCYRSPGEDQLSEFVYVGHLSGIRLOVKDKDETLVWELVYKPGDLMPATTLIPP 1320
Db 1261 KELELCYRSPGEDQLSEFVYVGHLSGIRLOVKDKDETLVWELVYKPGDLMPATTLIPP 1320
Qy 1321 ARTAVPSLPDAPOLLHFVDQYRQLARTVTSVEVLDKHLGQVLSQEOYERVLAEINTPS 1380
Db 1321 ARTAVPSLPDAPOLLHFVDQYRQLARTVTSVEVLDKHLGQVLSQEOYERVLAEINTPS 1380
Qy 1381 QMRKFLSLQSDWDRCKDGLYQALKETHPHILMELWEGSKGKGLPLSS 1429
Db 1381 QMRKFLSLQSDWDRCKDGLYQALKETHPHILMELWEGSKGKGLPLSS 1429

RESULT 3
US-10-028-374-15
; Sequence 15, Application US/10028374
; Publication No. US20030143706A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN LEUCINE-RICH REPEAT CONTAINING PROTEIN EXPRESSED
; TITLE OF INVENTION: PREDOMINATELY IN BONE MARROW, HLRRBM1

; FILE REFERENCE: D0067NP
; CURRENT APPLICATION NUMBER: US/10/028,374
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/257,773
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 1429
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-028-374-15

Query Match 100.0%; Score 7534; DB 12; Length 1429;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAGGAWGRACLVLEFLKKEELKEFOLLANKAHSRSSSGETPAQPEKTSGMVEASVLAQ 60
Db 1 MAGGAWGRACLVLEFLKKEELKEFOLLANKAHSRSSSGETPAQPEKTSGMVEASVLAQ 60
Qy 61 YGEQAWDLALHTWQMGRLSLCAQAQEGAGHSPSPYPSPHGLSPSQPTSTAVLMPW 120
Db 61 YGEQAWDLALHTWQMGRLSLCAQAQEGAGHSPSPYPSPHGLSPSQPTSTAVLMPW 120
Qy 121 IHLPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSPDHESPSQSPNAPTST 180
Db 121 IHLPAGCTQGSERRVLRQLPDTSGRRWREISASLLYQALPSPDHESPSQSPNAPTST 180
Qy 181 AVLGSWGPSPSLAPRQEAFTGOWPLDETSGIYYTEIREREREKSEKGRPPWAAVGT 240
Db 181 AVLGSWGPSPSLAPRQEAFTGOWPLDETSGIYYTEIREREREKSEKGRPPWAAVGT 240
Qy 241 PQOAMTSLQPHHPHPEPSVRESLCSWPMKNDENQKFTQLLLQRPSPRSDPLVKRSW 300
Db 241 PQOAMTSLQPHHPHPEPSVRESLCSWPMKNDENQKFTQLLLQRPSPRSDPLVKRSW 300
Qy 301 PDYVEENRGLHIEIRDLPFGGLDTQEPRIVLQGAAGIKSTLARQVKEAWGRGQLYGDR 360
Db 301 PDYVEENRGLHIEIRDLPFGGLDTQEPRIVLQGAAGIKSTLARQVKEAWGRGQLYGDR 360
Qy 361 FOHVYFSCRELAQSKVSLABELIKDGTATAPRIQLSRPERLLFILDGVDPEQWVLQ 420
Db 361 FOHVYFSCRELAQSKVSLABELIKDGTATAPRIQLSRPERLLFILDGVDPEQWVLQ 420
Qy 421 EPSSELCLHWSQPADALLGSLGKTLILPEASFILITARTALQNLIPSLQARWVEVLG 480
Db 421 EPSSELCLHWSQPADALLGSLGKTLILPEASFILITARTALQNLIPSLQARWVEVLG 480
Qy 481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNGKELWALCLVPWVSWLACTCLMQMKKEKL 540
Db 481 FSESSRKEYFYRYFTDERQAIRAFRLVKSNGKELWALCLVPWVSWLACTCLMQMKKEKL 540
Qy 541 TLTSKTTTTLCHYLAQAALQAQPLGQPLRDLCSLAAEGIWOKTFLFSPDDLRKHGIDGAI 600
Db 541 TLTSKTTTTLCHYLAQAALQAQPLGQPLRDLCSLAAEGIWOKTFLFSPDDLRKHGIDGAI 600
Qy 601 ISTFLKMGILQEHPIPLSYFHLFCQEFFAAMSVYVLEDEKGRGHSNCIIIDLEKTLAY 660
Db 601 ISTFLKMGILQEHPIPLSYFHLFCQEFFAAMSVYVLEDEKGRGHSNCIIIDLEKTLAY 660
Qy 661 GIHGLFGASTTRFLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLOPHSLESLSH 720
Db 661 GIHGLFGASTTRFLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLOPHSLESLSH 720
Qy 721 CLYETRNKFTLQVMAHFEMGCMVETDMELLVCTFCIKFSRHVKKLQIEGRQHRSTWS 780
Db 721 CLYETRNKFTLQVMAHFEMGCMVETDMELLVCTFCIKFSRHVKKLQIEGRQHRSTWS 780
Qy 781 PTMVVLFWRVPVTDAYWQILFSLVKVTRNLKELDLSGNSLSHSAVKSCLKTLRRPRCLLE 840
Db 781 PTMVVLFWRVPVTDAYWQILFSLVKVTRNLKELDLSGNSLSHSAVKSCLKTLRRPRCLLE 840

QY	841	TLRLAGCGLTAECDKDLAFGLRANQTLTDLDSFNVLTDAKAKHLQRLRQPSCKLRLQ	900
DB	841	TLRLAGCGLTAECDKDLAFGLRANQTLTDLDSFNVLTDAKAKHLQRLRQPSCKLRLQ	900
QY	901	LVS CGLTSCCCODLASVLASPSLKBELDLQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ	960
DB	901	LVS CGLTSCCCODLASVLASPSLKBELDLQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ	960
QY	961	TTLSDEMRQELRALAEQEKPOLLIIFSRKPSVMTPTTEGLDTGEMSNSTSLKQRLGSE	1020
DB	961	TTLSDEMRQELRALAEQEKPOLLIIFSRKPSVMTPTTEGLDTGEMSNSTSLKQRLGSE	1020
QY	1021	ASHVAQANLKLDDVSKIFPIAIEABESSPEVVPVELLCVPSAQDGLHTKPLGTDDEFW	1080
DB	1021	ASHVAQANLKLDDVSKIFPIAIEABESSPEVVPVELLCVPSAQDGLHTKPLGTDDEFW	1080
QY	1081	GPTGPVATEVWDKEKNLYRVHFPVAGSYRPNNTGLCFVNRREAVTVEIEFCVWDQFLGEIN	1140
DB	1081	GPTGPVATEVWDKEKNLYRVHFPVAGSYRPNNTGLCFVNRREAVTVEIEFCVWDQFLGEIN	1140
QY	1141	POHSMWVAGPLDIDIAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKBEGMLLEKPARV	1200
DB	1141	POHSMWVAGPLDIDIAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKBEGMLLEKPARV	1200
QY	1201	ELHHIVLENPSPGLGVLLKMHNLARFIPVTSVLLYHRVHPEEVTFLHYLIPSDCSIR	1260
DB	1201	ELHHIVLENPSPGLGVLLKMHNLARFIPVTSVLLYHRVHPEEVTFLHYLIPSDCSIR	1260
QY	1261	KELELCYRSPGEDQLFSEFYVGHLSGIRLQVKDKDETLVMEALVKPGDLMPATTLIPP	1320
DB	1261	KELELCYRSPGEDQLFSEFYVGHLSGIRLQVKDKDETLVMEALVKPGDLMPATTLIPP	1320
QY	1321	ARIAPSPDLADQLLHFVDQYREQLIARVTSVEVVLDKLHGQVLSQEOYERVLAE	1380
DB	1321	ARIAPSPDLADQLLHFVDQYREQLIARVTSVEVVLDKLHGQVLSQEOYERVLAE	1380
QY	1381	QMRKLFSLSQSDRCKDCGLYQALKETHPHLMELWEKSKKGLLPLSS	1429
DB	1381	QMRKLFSLSQSDRCKDCGLYQALKETHPHLMELWEKSKKGLLPLSS	1429

RESULT 4
US-10-183-770-15
; Sequence 15, Application US/10183770
; Publication No. US20030180812A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN LEUCINE-RICH REPEAT CONTAINING PROTEIN EXPRESSED
; FILE OF INVENTION: PREDOMINATELY IN BONE MARROW, HLRRB1
; FILE REFERENCE: D0067A CIP
; CURRENT APPLICATION NUMBER: US/10/183,770
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION NUMBER: US 60/257,773
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 10/028,374
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 15
; LENGTH: 1429
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-183-770-15

Query Match 100.0%; Score 7534; DB 12; Length 1429;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAGGAWGRACLYLEFLKKEELKEFQALLANKAHSRSSGSETPAQPKTSGMEVASYLVAQ	60
DB	1	MAGGAWGRACLYLEFLKKEELKEFQALLANKAHSRSSGSETPAQPKTSGMEVASYLVAQ	60
QY	61	YGEQRAWDLALHTWEQMLRSILCAQOEGAGHSPSPFPSPHLSGSPQSTSTAVLMPW	120

DB	61	YGEQRAWDLALHTWEQMLRSILCAQOEGAGHSPSPFPSPHLSGSPQSTSTAVLMPW	120
QY	121	THELPAGCTQCSERVLQPLDTSGRWRRETSASLLYQALPSSPDHSPSPESNPATST	180
DB	121	THELPAGCTQCSERVLQPLDTSGRWRRETSASLLYQALPSSPDHSPSPESNPATST	180
QY	181	AVLSGWSGPPPSLAPREQEAPGTOWPLDETSGIYYTIREREREKSKGRPPMAAVVGT	240
DB	181	AVLSGWSGPPPSLAPREQEAPGTOWPLDETSGIYYTIREREREKSKGRPPMAAVVGT	240
QY	241	PPQAHTSLQPHHPWPSPVRESLCTWPWKNEFNQKFTQLLLQRPHPRSQDPLVKRWS	300
DB	241	PPQAHTSLQPHHPWPSPVRESLCTWPWKNEFNQKFTQLLLQRPHPRSQDPLVKRWS	300
QY	301	PDYVEENHGHLIETRDLEFGPLDTPQEPRIVLQGAAGIGKSTLAPQVKEANGRGOLYDGR	360
DB	301	PDYVEENHGHLIETRDLEFGPLDTPQEPRIVLQGAAGIGKSTLAPQVKEANGRGOLYDGR	360
QY	361	FOHVFFYFSCRELAQSKVSVSLAELIGKGTATPAPIRQILSRPERLLFLDGVDEPGWLQ	420
DB	361	FOHVFFYFSCRELAQSKVSVSLAELIGKGTATPAPIRQILSRPERLLFLDGVDEPGWLQ	420
QY	421	EPSSCLCLHWSQPADALLGSLGKTLTILPEASFLITARTTALQNLIPSLBOARWVEVLG	480
DB	421	EPSSCLCLHWSQPADALLGSLGKTLTILPEASFLITARTTALQNLIPSLBOARWVEVLG	480
QY	481	FSESRKKEYFYRYFTDERQATRAFLVKSNKELWALCLVPMVSWLACTCLMQOMKREKL	540
DB	481	FSESRKKEYFYRYFTDERQATRAFLVKSNKELWALCLVPMVSWLACTCLMQOMKREKL	540
QY	541	TLTSTTTTTCLHYLAQALQAPLGPQPLRDCLSLAAGIWOKTFLSPDDLRKHGLDGA	600
DB	541	TLTSTTTTTCLHYLAQALQAPLGPQPLRDCLSLAAGIWOKTFLSPDDLRKHGLDGA	600
QY	601	ISTFLKMGILQEHPIPLSYFHLICQBFFAAMSYLEDEKRGKSHNCIIIDLEKTLBAY	660
DB	601	ISTFLKMGILQEHPIPLSYFHLICQBFFAAMSYLEDEKRGKSHNCIIIDLEKTLBAY	660
QY	661	GIHGLFGASTTRFLGLLSDEGEREMENI FHCRLSQGNLMQWVPSLQLLQPHSLES	720
DB	661	GIHGLFGASTTRFLGLLSDEGEREMENI FHCRLSQGNLMQWVPSLQLLQPHSLES	720
QY	721	CLYETRNKTFITQVMAHFEEMCMVETDMLLVCTFCIKFSRHVKKLQIIEGRQHRSTWS	780
DB	721	CLYETRNKTFITQVMAHFEEMCMVETDMLLVCTFCIKFSRHVKKLQIIEGRQHRSTWS	780
QY	781	PTMVVLFWRVPVTDAYWQILFSLVKVTRNLKELDLSGNSLSHSAVKSICKTLRRPRCLLE	840
DB	781	PTMVVLFWRVPVTDAYWQILFSLVKVTRNLKELDLSGNSLSHSAVKSICKTLRRPRCLLE	840
QY	841	TLRLAGCGLTAECDKDLAFGLRANQTLTDLDSFNVLTDAKAKHLQRLRQPSCKLRLQ	900
DB	841	TLRLAGCGLTAECDKDLAFGLRANQTLTDLDSFNVLTDAKAKHLQRLRQPSCKLRLQ	900
QY	901	LVS CGLTSDCCODLASVLASPSLKBELDLQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ	960
DB	901	LVS CGLTSDCCODLASVLASPSLKBELDLQNNLDDVGVRLLCEGLRHPACKLIRLGLDQ	960
QY	961	TTLSDEMRQELRALAEQEKPOLLIIFSRKPSVMTPTTEGLDTGEMSNSTSLKQRLGSE	1020
DB	961	TTLSDEMRQELRALAEQEKPOLLIIFSRKPSVMTPTTEGLDTGEMSNSTSLKQRLGSE	1020
QY	1021	ASHVAQANLKLDDVSKIFPIAIEABESSPEVVPVELLCVPSAQDGLHTKPLGTDDEFW	1080
DB	1021	ASHVAQANLKLDDVSKIFPIAIEABESSPEVVPVELLCVPSAQDGLHTKPLGTDDEFW	1080
QY	1081	GPTGPVATEVWDKEKNLYRVHFPVAGSYRPNNTGLCFVNRREAVTVEIEFCVWDQFLGEIN	1140
DB	1081	GPTGPVATEVWDKEKNLYRVHFPVAGSYRPNNTGLCFVNRREAVTVEIEFCVWDQFLGEIN	1140
QY	1141	POHSMWVAGPLDIDIAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFKBEGMLLEKPARV	1200

Db 1141 POHSMWVAGPLLDIAKAPGAVEAVHLPHPVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
Qy 1201 ELHHIVLENPSFPLGVLLKMIHNNALRFIPVTSVLLYHRVHPPEVTFHLYLIPSDCSIR 1260
Db 1201 ELHHIVLENPSFPLGVLLKMIHNNALRFIPVTSVLLYHRVHPPEVTFHLYLIPSDCSIR 1260
Qy 1261 KELELCYRSPGSDQLFSEFYVGHLSGIRLQVKDKDETLVWEALVKPGDLMPATTLIPP 1320
Db 1261 KELELCYRSPGSDQLFSEFYVGHLSGIRLQVKDKDETLVWEALVKPGDLMPATTLIPP 1320
Qy 1321 ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVDLKLHGQVLSOEQYERVLAEENTRPS 1380
Db 1321 ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVDLKLHGQVLSOEQYERVLAEENTRPS 1380
Qy 1381 QMRKLFSLSQSWDRCKCKDGLYQALKETHPHLMELWEKSGKGLPLSS 1429
Db 1381 QMRKLFSLSQSWDRCKCKDGLYQALKETHPHLMELWEKSGKGLPLSS 1429

RESULT 5
US-10-028-392-11
; Sequence 11, Application US/10028392
; Publication No. US20030087340A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN LEUCINE-RICH REPEAT CONTAINING PROTEIN EXPRESSED
; FILE REFERENCE: D0085.jp
; CURRENT APPLICATION NUMBER: US/10/028,392
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/259,479
; PRIOR FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: US 60/260,616
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 1429
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-028-392-11

Query Match 100.0%; Score 7534; DB 15; Length 1429;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAGGAWGRACLYLEFLKKEELKEFQLLANKAHSRSSGSETPAQPEKTSQMEVASYLVAQ 60
Db 1 MAGGAWGRACLYLEFLKKEELKEFQLLANKAHSRSSGSETPAQPEKTSQMEVASYLVAQ 60
Qy 61 YGEQRAWDALHTWQMGRLSLCAQAEAGAGHSPSPYSPHPHLSGSPSQPTSTAVLMPW 120
Db 61 YGEQRAWDALHTWQMGRLSLCAQAEAGAGHSPSPYSPHPHLSGSPSQPTSTAVLMPW 120
Qy 121 IHLEPAGCTQGSRRVRLQPLDTSGRWRREISASLLYQALPSPDHSPSQSNAPTST 180
Db 121 IHLEPAGCTQGSRRVRLQPLDTSGRWRREISASLLYQALPSPDHSPSQSNAPTST 180
Qy 181 AVLWGSGPPQSLAPREAGPTQWPLDTSIYITEIREREREKSGRPPAAVGT 240
Db 181 AVLWGSGPPQSLAPREAGPTQWPLDTSIYITEIREREREKSGRPPAAVGT 240
Qy 241 PQAHTSLQPHHHHPWSPRESLCSWPMKNEFDNQFTQLLLQPHRPSQDPLVKRSM 300
Db 241 PQAHTSLQPHHHHPWSPRESLCSWPMKNEFDNQFTQLLLQPHRPSQDPLVKRSM 300
Qy 301 PDVVEENRGLHIEIRDLFGPLDTPQPRIVILQGAAGIKSTILARQVKEAWGRGQLYGDR 360
Db 301 PDVVEENRGLHIEIRDLFGPLDTPQPRIVILQGAAGIKSTILARQVKEAWGRGQLYGDR 360
Qy 361 FQHVFFSCRELAQSKVSLAEILGKDGATATPAPIRQILSRPERLLFILDGVDPEGWLIQ 420
Db 361 FQHVFFSCRELAQSKVSLAEILGKDGATATPAPIRQILSRPERLLFILDGVDPEGWLIQ 420

RESULT 6

US-09-388-221-2

; Sequence 2, Application US/09388221A

Qy 421 EPSSELCHWVSQPADALLGSLGKTTILPEASFLITARTTALQNLIPSLQEARVVEVLG 480
Db 421 EPSSELCHWVSQPADALLGSLGKTTILPEASFLITARTTALQNLIPSLQEARVVEVLG 480
Qy 481 PSESRKKEYFYRYFTDERQAIRAPLAKSNKELWALCLVPVWSWLACTCLMOQMKRKEKL 540
Db 481 PSESRKKEYFYRYFTDERQAIRAPLAKSNKELWALCLVPVWSWLACTCLMOQMKRKEKL 540
Qy 541 TLTSKTTTTLCHLYLAQALQAPLQPOLRDICSLAAEGIWQKTLFSPDDLRKHGLDGA 600
Db 541 TLTSKTTTTLCHLYLAQALQAPLQPOLRDICSLAAEGIWQKTLFSPDDLRKHGLDGA 600
Qy 601 ISTFLKMGILQEHPIPLSYFHLFCQFFFAAMSYYLEDEKGRGKHSNCIIIDLEKTL 660
Db 601 ISTFLKMGILQEHPIPLSYFHLFCQFFFAAMSYYLEDEKGRGKHSNCIIIDLEKTL 660
Qy 661 GIHGLFGASTTRFLGILLSDGEREMENI FHCRLSQGRNLMQWVPSLQILLQPHSLES 720
Db 661 GIHGLFGASTTRFLGILLSDGEREMENI FHCRLSQGRNLMQWVPSLQILLQPHSLES 720
Qy 721 CLYETRNKTFITQVMAHFEEMGMCVETDMLLVCTFCIKFSRHVKKQLIEGRQHRSTWS 780
Db 721 CLYETRNKTFITQVMAHFEEMGMCVETDMLLVCTFCIKFSRHVKKQLIEGRQHRSTWS 780
Qy 781 PTMVVLFWRVPVTDAYWQILFSLVKVTRNLKELDLSGNSLSHSAVKSCLKTLRRPRCLLE 840
Db 781 PTMVVLFWRVPVTDAYWQILFSLVKVTRNLKELDLSGNSLSHSAVKSCLKTLRRPRCLLE 840
Qy 841 TURLAGCGLTABDCXDLAFGLRANQTLTDLDSFNVLTDAGAKHLQRLQRBQSKLQRLQ 900
Db 841 TURLAGCGLTABDCXDLAFGLRANQTLTDLDSFNVLTDAGAKHLQRLQRBQSKLQRLQ 900
Qy 901 LVSCGLTSDCCODLASVLSASPLKELDLQNNLDDVGVRLICEGLRHHPACKLIRGLDQ 960
Db 901 LVSCGLTSDCCODLASVLSASPLKELDLQNNLDDVGVRLICEGLRHHPACKLIRGLDQ 960
Qy 961 TTLSDEMREQLRALQEQKPOLLI FSRKRP SVMTPTTEGLDTGEMSNSTSLKQRLGSE 1020
Db 961 TTLSDEMREQLRALQEQKPOLLI FSRKRP SVMTPTTEGLDTGEMSNSTSLKQRLGSE 1020
Qy 1021 ASHVAQANLKLLDVSKIPIPIAIEABESSPEVVPVELLCVPSPASQGLHTKPLGTTDDDPW 1080
Db 1021 ASHVAQANLKLLDVSKIPIPIAIEABESSPEVVPVELLCVPSPASQGLHTKPLGTTDDDPW 1080
Qy 1081 GPTGPVATEVVDKKNLYRVHPFVAGSYRWPNPTGLCFVWREAVTVEIEFCVWDQFLGEIN 1140
Db 1081 GPTGPVATEVVDKKNLYRVHPFVAGSYRWPNPTGLCFVWREAVTVEIEFCVWDQFLGEIN 1140
Qy 1141 PQHSMWVAGPLLDIAKAPGAVEAVHLPHPVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
Db 1141 PQHSMWVAGPLLDIAKAPGAVEAVHLPHPVALQGGHVDTSLFQMAHFKEEGMLLEKPARV 1200
Qy 1201 ELHHIVLENPSFPLGVLLKMIHNNALRFIPVTSVLLYHRVHPPEVTFHLYLIPSDCSIR 1260
Db 1201 ELHHIVLENPSFPLGVLLKMIHNNALRFIPVTSVLLYHRVHPPEVTFHLYLIPSDCSIR 1260
Qy 1261 KELELCYRSPGSDQLFSEFYVGHLSGIRLQVKDKDETLVWEALVKPGDLMPATTLIPP 1320
Db 1261 KELELCYRSPGSDQLFSEFYVGHLSGIRLQVKDKDETLVWEALVKPGDLMPATTLIPP 1320
Qy 1321 ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVDLKLHGQVLSOEQYERVLAEENTRPS 1380
Db 1321 ARIAVPSPLDAPQLLHFVDQYREQLIARVTSVEVVDLKLHGQVLSOEQYERVLAEENTRPS 1380
Qy 1381 QMRKLFSLSQSWDRCKCKDGLYQALKETHPHLMELWEKSGKGLPLSS 1429
Db 1381 QMRKLFSLSQSWDRCKCKDGLYQALKETHPHLMELWEKSGKGLPLSS 1429

Publication No. US20020192643A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. US20020192643A1el Card Proteins Involved in Cell Death Regul
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388, 221A
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1473
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-388-221-2

Query Match 99.6%; Score 7502; DB 10; Length 1473;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 1429; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

QY 1 MAGGAWRLACYLEFLKKEELKEFQLLANKAHSRSSGETPAQPEKTSMEVASYLVAQ 60
DB 1 MAGGAWRLACYLEFLKKEELKEFQLLANKAHSRSSGETPAQPEKTSMEVASYLVAQ 60
QY 61 YGEQRAWDALHTWEQMLRSLCAQAEQAGHSPSPFPYSPSEPHLGSPQPTSTAVLMPW 120
DB 61 YGEQRAWDALHTWEQMLRSLCAQAEQAGHSPSPFPYSPSEPHLGSPQPTSTAVLMPW 120
QY 121 IHELPAGTCQGSRRVRLQLPDTSGRWRREISASLLYQALPSSPDHESPSQSPNAPTST 180
DB 121 IHELPAGTCQGSRRVRLQLPDTSGRWRREISASLLYQALPSSPDHESPSQSPNAPTST 180
QY 181 AVLGSNGSPPOPSLAPREQAPCTQPLDETSGIYVTEIRERERESEKGRPPWAAVGT 240
DB 181 AVLGSNGSPPOPSLAPREQAPCTQPLDETSGIYVTEIRERERESEKGRPPWAAVGT 240
QY 241 PPOAHTSLOPHHHHPSPVRESLCSTWPKNEDFNQFTQLLLQRPHPRSQDPLVKRSW 300
DB 241 PPOAHTSLOPHHHHPSPVRESLCSTWPKNEDFNQFTQLLLQRPHPRSQDPLVKRSW 300
QY 301 PDYVEENRGHLIBIRLFGPLDTPRIVILQGAAGIKSTLARQVKEAWRGQLYGR 360
DB 301 PDYVEENRGHLIBIRLFGPLDTPRIVILQGAAGIKSTLARQVKEAWRGQLYGR 360
QY 361 FOHVYFSCRELAQSVSLAEILGKGTATPAPIQILSRPERLLFILDGVDPEGWLQ 420
DB 361 FOHVYFSCRELAQSVSLAEILGKGTATPAPIQILSRPERLLFILDGVDPEGWLQ 420
QY 421 EPSSELCHWSQPQADALIGSLGKLTILPEASFLITARTALQNLIPSEQARWVEVLG 480
DB 421 EPSSELCHWSQPQADALIGSLGKLTILPEASFLITARTALQNLIPSEQARWVEVLG 480
QY 481 FSESSRKEYFYRTDERQAIRAFRLVKNKELWALCLVPWVSWLACTCLMQMKREKL 540
DB 481 FSESSRKEYFYRTDERQAIRAFRLVKNKELWALCLVPWVSWLACTCLMQMKREKL 540
QY 541 TLTSKTTTTLCHYLQAQAQPLGQPLDRDCLSLAEGIWOKTLPSPDRLKHGJLDGAI 600
DB 541 TLTSKTTTTLCHYLQAQAQPLGQPLDRDCLSLAEGIWOKTLPSPDRLKHGJLDGAI 600
QY 601 ISTFLKMGILQEHPIPLSYFHLCPQEPFAAMSVYVLEDEKGRGHSNCIIDLKTELEY 660
DB 601 ISTFLKMGILQEHPIPLSYFHLCPQEPFAAMSVYVLEDEKGRGHSNCIIDLKTELEY 660
QY 661 GIHGLFGASTTRPLGLLSDGGEREMENIFHCLRSQGRNLMQWVPSLQLLQPHSLESJH 720
DB 661 GIHGLFGASTTRPLGLLSDGGEREMENIFHCLRSQGRNLMQWVPSLQLLQPHSLESJH 720
QY 721 CLYETENKFTLVQMAHFEEMGCVETDMLLVCCTCIKESFHVKKQLIEGQHRSTWS 780
DB 721 CLYETENKFTLVQMAHFEEMGCVETDMLLVCCTCIKESFHVKKQLIEGQHRSTWS 780
QY 781 PTMVVLFWRVPVTDAYWQILFVSVLKVTNRNLKELDLSGNSLSHSAVKSCKTLRRRRCLE 840

DB 781 PTMVVLFWRVPVTDAYWQILFVSVLKVTNRNLKELDLSGNSLSHSAVKSCKTLRRRRCLE 840
QY 841 TURLAGCGLTABDCXDLAFGLRANOTLTDELDFSNVLTADAGAKHLQRLRQPSCKLRLQ 900
DB 841 TURLAGCGLTABDCXDLAFGLRANOTLTDELDFSNVLTADAGAKHLQRLRQPSCKLRLQ 900
QY 901 LVSCGLTSDCCODLASVLSASPSLKELDLQNNLDDVGVRLLCCEGLRHCPACKLIRGLDQ 960
DB 901 LVSCGLTSDCCODLASVLSASPSLKELDLQNNLDDVGVRLLCCEGLRHCPACKLIRGLDQ 960
QY 961 TTLSDEMRQELRALRQEKPOLLI FSRKPSVMTPTGDLDTGEMSNSTSLKQRGSGERA 1020
DB 961 TTLSDEMRQELRALRQEKPOLLI FSRKPSVMTPTGDLDTGEMSNSTSLKQRGSGERA 1020
QY 1021 ASHVAQANLLLDVSKIPPIAEIAEESSEVPVVELLCVPSPASQGDHLTKPLGTDDDFW 1080
DB 1021 ASHVAQANLLLDVSKIPPIAEIAEESSEVPVVELLCVPSPASQGDHLTKPLGTDDDFW 1080
QY 1081 GPTGPVATEVVDKEKNLYRVHFPFVAGSYRWPNTGLCFVWREAVTVIEIFCVWDQFLGEIN 1140
DB 1081 GPTGPVATEVVDKEKNLYRVHFPFVAGSYRWPNTGLCFVWREAVTVIEIFCVWDQFLGEIN 1140
QY 1141 PQSHMWVAGPLLDI IKAEPGAVEAHLPHFVALQGHVDTSLFQMAHFKKEGMLLEKPARV 1200
DB 1141 PQSHMWVAGPLLDI IKAEPGAVEAHLPHFVALQGHVDTSLFQMAHFKKEGMLLEKPARV 1200
QY 1201 ELHHIVLENPSPLGVLLKMIHNALRFPVTSVLLVHRVHPEVTHLYLIPSDCSIR 1260
DB 1201 ELHHIVLENPSPLGVLLKMIHNALRFPVTSVLLVHRVHPEVTHLYLIPSDCSIR 1260
QY 1261 -----KELELCYRSPGEDQLF 1276
DB 1261 KAIDDLKMKFQVRTHKPPPLTPLYMGCRYTVSGSGSGMLEILPKLELCYRSPGEDQLF 1320
QY 1277 SEFYVGHGSGIRLQVKKOETLWELVLPGLDMPATTLIPPARIAVPSPLDAPQLLH 1336
DB 1321 SEFYVGHGSGIRLQVKKOETLWELVLPGLDMPATTLIPPARIAVPSPLDAPQLLH 1380
QY 1337 FVDQYREQLIARVTSVEVVDLKHQVLSQEQYERVLAEINTRPSQMRKLFSLSQSWDRKC 1396
DB 1381 FVDQYREQLIARVTSVEVVDLKHQVLSQEQYERVLAEINTRPSQMRKLFSLSQSWDRKC 1440
QY 1397 KDGLYQALKETHPHLIMELWEKSGKGLPLSS 1429
DB 1441 KDGLYQALKETHPHLIMELWEKSGKGLPLSS 1473

RESULT 7
US-10-028-374-3
; Sequence 3, Application US/10028374
; Publication No. US20030143706A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN LEUCINE-RICH REPEAT CONTAINING PROTEIN EXPRESSED
; FILE OF INVENTION: PREDOMINATELY IN BONE MARROW, HLRRB1
; FILE REFERENCE: D0067NP
; CURRENT APPLICATION NUMBER: US/10/028,374
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/257,773
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 1429
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-028-374-3

Query Match 99.4%; Score 7488; DB 12; Length 1429;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1420; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAGGAWRLACYLEFLKKEELKEFQLLANKAHSRSSGETPAQPEKTSMEVASYLVAQ 60

Db	1	MAGGAWRLACYLEFLKKEELKEFOLLANKAHSRSSGETPAQPEKTSQMEVASYLVAQ	60
QY	61	YGEQRAWDALHTWEQWLRLSCAQAEAGAGHSPSPFPSEPHLGSPQPTSTAVLMPW	120
Db	61	YGEQRAWDALHTWEQWLRLSCAQAEAGAGHSPSPFPSEPHLGSPQPTSTAVLMPW	120
QY	121	IHELPAQCTQGSRRVRLQPLDTSGRRWREISASLLYQALPSPDHESPSQESPNAPTST	180
Db	121	IHELPAQCTQGSRRVRLQPLDTSGRRWREISASLLYQALPSPDHESPSQESPNAPTST	180
QY	181	AVLGSGSPPOPSLAHQEAPCTQWPLDETSGIYYTEIRERERESEKGRPPAAVVGCT	240
Db	181	AVLGSGSPPOPSLAHQEAPCTQWPLDETSGIYYTEIRERERESEKGRPPAAVVGCT	240
QY	241	PPQAHSLQPHHPHWPSPVRESLCSWPWKNEDFNOKFTQLLLQRPSPRSDPLVKRSW	300
Db	241	PPQAHSLQPHHPHWPSPVRESLCSWPWKNEDFNOKFTQLLLQRPSPRSDPLVKRSW	300
QY	301	PDYVEENRGLHIEIRDLFGGLDTPRIVILQGAAGIGKSTLARQVKEAWRGQLYGDR	360
Db	301	PDYVEENRGLHIEIRDLFGGLDTPRIVILQGAAGIGKSTLARQVKEAWRGQLYGDR	360
QY	361	FQHVFPSCRELAQSVSLAEILI GKGDTATPAPIRQILSRPERLILFDGVDPEQWVLQ	420
Db	361	FQHVFPSCRELAQSVSLAEILI GKGDTATPAPIRQILSRPERLILFDGVDPEQWVLQ	420
QY	421	EPSSSELCHWSQOPDADLLGSLGKTIILPEASFLITARTALONLIPSLQEARWVEVLG	480
Db	421	EPSSSELCHWSQOPDADLLGSLGKTIILPEASFLITARTALONLIPSLQEARWVEVLG	480
QY	481	FSSESRKEYFYRTDERQAIRFLVKSINKELWALCLVPWVSWLACTCLMQQMKRKEKL	540
Db	481	FSSESRKEYFYRTDERQAIRFLVKSINKELWALCLVPWVSWLACTCLMQQMKRKEKL	540
QY	541	TLTSKTTTTLCHLYLAQAQPLGQPLRDLCSLAAGIWKTKTLFSPDDLRKHGLDGA	600
Db	541	TLTSKTTTTLCHLYLAQAQPLGQPLRDLCSLAAGIWKTKTLFSPDDLRKHGLDGA	600
QY	601	ISTFLKMGILQEHPIPLSYFIHLFCQEPFAAMSYYLEDEKGRGKHSNCIIDLKLEAY	660
Db	601	ISTFLKMGILQEHPIPLSYFIHLFCQEPFAAMSYYLEDEKGRGKHSNCIIDLKLEAY	660
QY	661	GIHLFGCASTRFLGLLSDEGEREMENIFCHRLSQGNLMQWVPSLQLLLOPHSLES	720
Db	661	GIHLFGCASTRFLGLLSDEGEREMENIFCHRLSQGNLMQWVPSLQLLLOPHSLES	720
QY	721	CLYETRKNKTPLTQMAHFEEMGMCVETDMELLVCTFCIKFSPRHVKKQLIEGRQHR	780
Db	721	CLYETRKNKTPLTQMAHFEEMGMCVETDMELLVCTFCIKFSPRHVKKQLIEGRQHR	780
QY	781	PTMVVLFRWVPVTDAYWQILFVSVLKYTRNLKELDLSGNSLSHSAVKSCLKTLRRPCL	840
Db	781	PTMVVLFRWVPVTDAYWQILFVSVLKYTRNLKELDLSGNSLSHSAVKSCLKTLRRPCL	840
QY	841	TLRLAGCGLTAEDCKDLAFLGRANQTLTDLDSFNVLTDAGAKHLQRLRQPSCKLQRLQ	900
Db	841	TLRLAGCGLTAEDCKDLAFLGRANQTLTDLDSFNVLTDAGAKHLQRLRQPSCKLQRLQ	900
QY	901	LVSCGLTSDCCQDLASVLSASPSLKELDLQNNLDDVGVRLLCEGRHPACKLIRLGLDQ	960
Db	901	LVSCGLTSDCCQDLASVLSASPSLKELDLQNNLDDVGVRLLCEGRHPACKLIRLGLDQ	960
QY	961	TTLSDEMROELRALEQKPOLLIIFSRKPSVMTPTGLDTGENSNSTSSLKRORLGSE	1020
Db	961	TTLSDEMROELRALEQKPOLLIIFSRKPSVMTPTGLDTGENSNSTSSLKRORLGSE	1020
QY	1021	ASHVAQANLKLDDVSKIFPTAEIAESSPEVVPVELLCVPSPASQGLDHTKPLGTDDDFW	1080
Db	1021	ASHVAQANLKLDDVSKIFPTAEIAESSPEVVPVELLCVPSPASQGLDHTKPLGTDDDFW	1080
QY	1081	GPTGPVATEVVDKKNLYRVHFPVAGSYRWPNGLCFVVRVREAVTVEIEFCVNDQFLGEIN	1140

Db	1081	GPTGPVATEVVDKKNLYRVHFPVAGSYRWPNGLCFVVRVREAVTVEIEFCVNDQFLGEIN	1140
QY	1141	POHSVMWAGPLLDIKAEPGAVEAVHLPHFEVALQGGHVDTSLFQMAHFKEEGMLLEKPARV	1200
Db	1141	POHSVMWAGPLLDIKAEPGAVEAVHLPHFEVALQGGHVDTSLFQMAHFKEEGMLLEKPARV	1200
QY	1201	ELHHIVLENPSPSLGVLKMIHNNALRFIPVTSVLLYHRVHPEVTFHLYLIPSDCSIR	1260
Db	1201	ELHHIVLENPSPSLGVLKMIHNNALRFIPVTSVLLYHRVHPEVTFHLYLIPSDCSIR	1260
QY	1261	KELELCYSPGSDQLFSEFYVGHLSGIRLOVKDKKDETLVWEALVKPGDLMPATTLIPP	1320
Db	1261	KELELCYSPGSDQLFSEFYVGHLSGIRLOVKDKKDETLVWEALVKPGDLMPATTLIPP	1320
QY	1321	ARIAYPSPLDAPQLLHFVDQYREQLIARTSVSEVVLDKLHGQVLSQEQYERVLAE	1380
Db	1321	ARIAYPSPLDAPQLLHFVDQYREQLIARTSVSEVVLDKLHGQVLSQEQYERVLAE	1380
QY	1381	QMRKLFSLSQSWDRCKDGLYQALKETHPHILMELWEKSGKGLPLSS	1429
Db	1381	QMRKLFSLSQSWDRCKDGLYQALKETHPHILMELWEKSGKGLPLSS	1429

RESULT 8
US-10-183-770-3
; Sequence 3, Application US/10183770
; Publication No. US20030180812A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: A NOVEL HUMAN LEUCINE-RICH REPEAT CONTAINING PROTEIN EXPRESSED
; FILE REFERENCE: D0067A CIP
; CURRENT FILING DATE: 2002-06-27
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 60/257,773
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1429
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-183-770-3

Query Match 99.4%; Score 7488; DB 12; Length 1429;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1420; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY	1	MAGGAWRLACYLEFLKKEELKEFOLLANKAHSRSSGETPAQPEKTSQMEVASYLVAQ	60
Db	1	MAGGAWRLACYLEFLKKEELKEFOLLANKAHSRSSGETPAQPEKTSQMEVASYLVAQ	60
QY	61	YGEQRAWDALHTWEQWLRLSCAQAEAGAGHSPSPFPSEPHLGSPQPTSTAVLMPW	120
Db	61	YGEQRAWDALHTWEQWLRLSCAQAEAGAGHSPSPFPSEPHLGSPQPTSTAVLMPW	120
QY	121	IHELPAQCTQGSRRVRLQPLDTSGRRWREISASLLYQALPSPDHESPSQESPNAPTST	180
Db	121	IHELPAQCTQGSRRVRLQPLDTSGRRWREISASLLYQALPSPDHESPSQESPNAPTST	180
QY	181	AVLGSGSPPOPSLAHQEAPCTQWPLDETSGIYYTEIRERERESEKGRPPAAVVGCT	240
Db	181	AVLGSGSPPOPSLAHQEAPCTQWPLDETSGIYYTEIRERERESEKGRPPAAVVGCT	240
QY	241	PPQAHSLQPHHPHWPSPVRESLCSWPWKNEDFNOKFTQLLLQRPSPRSDPLVKRSW	300
Db	241	PPQAHSLQPHHPHWPSPVRESLCSWPWKNEDFNOKFTQLLLQRPSPRSDPLVKRSW	300
QY	301	PDYVEENRGLHIEIRDLFGGLDTPRIVILQGAAGIGKSTLARQVKEAWRGQLYGDR	360
Db	301	PDYVEENRGLHIEIRDLFGGLDTPRIVILQGAAGIGKSTLARQVKEAWRGQLYGDR	360

361 FOHVYFSCRELAQSKVSLAEILGKDGATATPAPIRQILSRPERLLFILDGVDPEGVWLQ 420
Db
361 FOHVYFSCRELAQSKVSLAEILGKDGATATPAPIRQILSRPERLLFILDGVDPEGVWLQ 420
Qy 421 EPSSELCHWSQPPADALLGSLGKTLTILPEASFLITARTTALQNLIPSLQARWVEVLG 480
Db 421 EPSSELCHWSQPPADALLGSLGKTLTILPEASFLITARTTALQNLIPSLQARWVEVLG 480
Qy 481 FSESSRKEFYRYFTDERQAIRAFLVKSNKELWALCLVPWVSWLACTCLMOQMKRKEKL 540
Db 481 FSESSRKEFYRYFTDERQAIRAFLVKSNKELWALCLVPWVSWLACTCLMOQMKRKEKL 540
Qy 541 TLTSKTTTTLCHLYLAQALQALQPLGQPLRDLCSLAAEGIWQKTLFSPDDLRKHGLDGA 600
Db 541 TLTSKTTTTLCHLYLAQALQALQPLGQPLRDLCSLAAEGIWQKTLFSPDDLRKHGLDGA 600
Qy 601 ISTFLKMGILQEHPIPLSYSFHLCQFEFFAAMSVYLEDEKGRGKHSNCIIDLKTLLEY 660
Db 601 ISTFLKMGILQEHPIPLSYSFHLCQFEFFAAMSVYLEDEKGRGKHSNCIIDLKTLLEY 660
Qy 661 GIHGLFGASTTRFLGLLSDGEREMENIFHCRLSQGNLMQWVPSLQLLLOPHSLESILH 720
Db 661 GIHGLFGASTTRFLGLLSDGEREMENIFHCRLSQGNLMQWVPSLQLLLOPHSLESILH 720
Qy 721 CLYETRNTKTLTQVMAHFEEMGCMVETDMLLVCFCIKFSRHVKQLQIEGRQHRSTWS 780
Db 721 CLYETRNTKTLTQVMAHFEEMGCMVETDMLLVCFCIKFSRHVKQLQIEGRQHRSTWS 780
Qy 781 PTMVVLFRWPVPTDAYWQILFVSLKVTNRNLKELDLGNSLSHSAVSKLCTLRPRCLLE 840
Db 781 PSMVVLFRWPVPTDAYWQILFVSLKVTNRNLKELDLGNSLSHSAVSKLCTLRPRCLLE 840
Qy 841 TRLACGGLTAEDCKOLAFGLRANOTLTLDLSFNVLTDAGAKHLQRLRQSPCKLQRLQ 900
Db 841 TRLACGGLTAEDCKOLAFGLRANOTLTLDLSFNVLMDAGAKHLQRLRQSPCKLQRLQ 900
Qy 901 LVSCGTLSDCCODLASVLSASPSLKELDLQNNLDVGVRLLCGRHHPACKLIRLGLDQ 960
Db 901 LVSCGTLSDCCODLASVLSASPSLKELDLQNNLDVGVRLLCGRHHPACKLIRLGLDQ 960
Qy 961 TTLSDEMROELRALEQEKQLLIFSRKPSVMTPTBGLDTGEMSNSTSLKQRQLGSE 1020
Db 961 TTLSDEMROELRALEQEKQLLIFSRKPSVMTPIEGLDTGEMSNSTSLKQRQLGSE 1020
Qy 1021 ASHVAQNLKLDVSKIPPIAEIAESSSEVPVVELLCVPSPASQGLHTKPLGTDDDFW 1080
Db 1021 ASHVAQNLKLDVSKIPPIAEIAESSSEVPVVELLCVPSPASQGLHTKPLGTDDDFW 1080
Qy 1081 GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVMEATVVEIEFCVMDQFLGEIN 1140
Db 1081 GPTGPVATEVVDKEKNLYRVHFPVAGSYRWPNTGLCFVMEATVVEIEFCVMDQFLGEIN 1140
Qy 1141 PQHSMWVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFEKGMLLEKPARV 1200
Db 1141 PQHSMWVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFEKGMLLEKPARV 1200
Qy 1201 ELHHIVLENPSPGLGVLLKMHINARLFTPVTSVLLYHRVPEEVTFLHYLIPSDCSR 1260
Db 1201 ELHHIVLENPSPGLGVLLKMHINARLFTPVTSVLLYHRVPEEVTFLHYLIPSDCSR 1260
Qy 1261 KELELCYRSGEDOLFSEFVVGHLGSGIRLOVKDKDETLVMEALVKPGDLMPATTLIPP 1320
Db 1261 KELELCYRSGEDOLFSEFVVGHLGSGIRLOVKDKDETLVMEALVKPGDLMPATTLIPP 1320
Qy 1321 ARTAVPSPLDAPOLLHFVQYRQLIARTVTSVEVLDKLHGQVLSQEQYERVLAEINTPS 1380
Db 1321 ACTAVPSPLDAPOLLHFVQYRQLIARTVTSVEVLDKLHGQVLSQEQYERVLAEINTPS 1380
Qy 1381 QMRKLFSLSQSWDRKCKDGLYQALKETHPHILMELWEKSGKGLPLSS 1429
Db 1381 QMRKLFSLSQSWDRKCKDGLYQALKETHPHILMELWEKSGKGLPLSS 1429

RESULT 9
US-09-388-221-4
; Sequence 4, Application US/09388221A
; Publication NO. US20020192843A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. US20020192843A1el Card Proteins Involved in Cell Death Regula
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388, 221A
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1399
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-388-221-4

Query Match 97.7%; Score 7364; DB 10; Length 1399;
Best Local Similarity 97.9%; Pred No. 0;
Matches 1399; Conservative 0; Mismatches 0; Indels 30; Gaps 1;
Qy 1 MAGGAWGLACYLEFLKKEELKEFQLLANKAHSRSSGETPAQPEKTSMEVASYLVAQ 60
Db 1 MAGGAWGLACYLEFLKKEELKEFQLLANKAHSRSSGETPAQPEKTSMEVASYLVAQ 60
Qy 61 YGEQAWDLALHTWQMGRLSLCAQAEAGHSPSPFPSPSEPHLGSPQSTSTAVLMPW 120
Db 61 YGEQAWDLALHTWQMGRLSLCAQAEAGHSPSPFPSPSEPHLGSPQSTSTAVLMPW 120
Qy 121 HELPAGCTQGSERRVLQPLDTSRRWRREISASLLYQALPSSPDHESPSQSPNAPTST 180
Db 121 HELPAGCTQGSERRVLQPLDTSRRWRREISASLLYQALPSSPDHESPSQSPNAPTST 180
Qy 181 AVLGWGWSPPSLAPREQEAPGTQWPLDTSIGIYYTEIREREREKSEKGRPPAAVVG 240
Db 181 AVLGWGWSPPSLAPREQEAPGTQWPLDTSIGIYYTEIREREREKSEKGRPPAAVVG 240
Qy 241 PQAHTSLQPHHPHWPSPVRESLCSWTWPKNEDFNQKTQLLLQRPSPRQDPLVKRSW 300
Db 241 PQAHTSLQPHHPHWPSPVRESLCSWTWPKNEDFNQKTQLLLQRPSPRQDPLVKRSW 300
Qy 301 PDYVENRGLHLEIRDLFGPLDTCPRIVILQGAAGICKSTARQVKEAWGRGQLYGR 360
Db 301 PDYVENRGLHLEIRDLFGPLDTCPRIVILQGAAGICKSTARQVKEAWGRGQLYGR 360
Qy 361 FOHVYFSCRELAQSKVSLAEILGKDGATATPAPIRQILSRPERLLFILDGVDPEGVWLQ 420
Db 361 FOHVYFSCRELAQSKVSLAEILGKDGATATPAPIRQILSRPERLLFILDGVDPEGVWLQ 420
Qy 421 EPSSELCHWSQPPADALLGSLGKTLTILPEASFLITARTTALQNLIPSLQARWVEVLG 480
Db 421 EPSSELCHWSQPPADALLGSLGKTLTILPEASFLITARTTALQNLIPSLQARWVEVLG 480
Qy 481 FSESSRKEFYRYFTDERQAIRAFLVKSNKELWALCLVPWVSWLACTCLMOQMKRKEKL 540
Db 481 FSESSRKEFYRYFTDERQAIRAFLVKSNKELWALCLVPWVSWLACTCLMOQMKRKEKL 540
Qy 541 TLTSKTTTTLCHLYLAQALQALQPLGQPLRDLCSLAAEGIWQKTLFSPDDLRKHGLDGA 600
Db 541 TLTSKTTTTLCHLYLAQALQALQPLGQPLRDLCSLAAEGIWQKTLFSPDDLRKHGLDGA 600
Qy 601 ISTFLKMGILQEHPIPLSYSFHLCQFEFFAAMSVYLEDEKGRGKHSNCIIDLKTLLEY 660
Db 601 ISTFLKMGILQEHPIPLSYSFHLCQFEFFAAMSVYLEDEKGRGKHSNCIIDLKTLLEY 660
Qy 661 GIHGLFGASTTRFLGLLSDGEREMENIFHCRLSQGNLMQWVPSLQLLLOPHSLESILH 720
Db 661 GIHGLFGASTTRFLGLLSDGEREMENIFHCRLSQGNLMQWVPSLQLLLOPHSLESILH 720
Qy 721 CLYETRNTKTLTQVMAHFEEMGCMVETDMLLVCFCIKFSRHVKQLQIEGRQHRSTWS 780
Db 721 CLYETRNTKTLTQVMAHFEEMGCMVETDMLLVCFCIKFSRHVKQLQIEGRQHRSTWS 780

Db	721	CLYETRNKFTLTQVMAHFEEMGCVETDMLLVCTFCIKFSRHVKQLQLIEGRQHRSTWS	780
Qy	781	PTMVLFWRVPTVDAYWQILFSLVKVTRNLKELDLGNSLSHSAVKSCKTLRRPRCLLE	840
Db	781	PTMVLFWRVPTVDAYWQILFSLVKVTRNLKELDLGNSLSHSAVKSCKTLRRPRCLLE	840
Qy	841	TURLAGCGLTABDCCKDLAFGLRANOTLTLDLSFNVLTDAGAKHLCORLQPSCKLQRLQ	900
Db	841	TURLAGCGLTABDCCKDLAFGLRANOTLTLDLSFNVLTDAGAKHLCORLQPSCKLQRLQ	900
Qy	901	LVSCLGTSDDCCDLASVLSASPSLKDLDQNNLDVGVRLLCCEGLRHLPACKLIRLGLDQ	960
Db	901	LVSCLGTSDDCCDLASVLSASPSLKDLDQNNLDVGVRLLCCEGLRHLPACKLIRLGLDQ	960
Qy	961	TTLSDMRQELRALEQEKPOLLIIFSRKFPVMTPTTEGLDTGEMSNSTSLKQRLGSERA	1020
Db	958	-----KFSVMTPTTEGLDTGEMSNSTSLKQRLGSERA	990
Qy	1021	ASHVAQANLKLDVSKIPPIAIEBESSPEVVPVELLCVPSASQGDHLTKPLGTTDDDFW	1080
Db	991	ASHVAQANLKLDVSKIPPIAIEBESSPEVVPVELLCVPSASQGDHLTKPLGTTDDDFW	1050
Qy	1081	GPTGPVATEVVDKKNLYRVHPFVAGSVRWPNVTGLCFVNRVAVTVEIEFCVWDQFLGEIN	1140
Db	1051	GPTGPVATEVVDKKNLYRVHPFVAGSVRWPNVTGLCFVNRVAVTVEIEFCVWDQFLGEIN	1110
Qy	1141	POHSMWVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFEKGMLEKPARV	1200
Db	1111	POHSMWVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFEKGMLEKPARV	1170
Qy	1201	ELHHVLENPSPGLVLLKMIHNALRFTPVTSVLLYHRVHPVEVTHLYLIPSDCSIR	1260
Db	1171	ELHHVLENPSPGLVLLKMIHNALRFTPVTSVLLYHRVHPVEVTHLYLIPSDCSIR	1230
Qy	1261	KELELCYRSPGDDOLFSEFVYVGHGSGIRLOVKDKOETLVWEALVKPGDLMPATLIPP	1320
Db	1231	KELELCYRSPGDDOLFSEFVYVGHGSGIRLOVKDKOETLVWEALVKPGDLMPATLIPP	1290
Qy	1321	ARIAPVSPDLAQQLLHFVDQYREQLIARTSVSEVLDKLGHVLSQEOQYERVLAEINTRPS	1380
Db	1291	ARIAPVSPDLAQQLLHFVDQYREQLIARTSVSEVLDKLGHVLSQEOQYERVLAEINTRPS	1350
Qy	1381	QMRKLFSLQSQWDRKCKDGLYQALKETHPHLIMELWEKSKGLIPLSS	1429
Db	1351	QMRKLFSLQSQWDRKCKDGLYQALKETHPHLIMELWEKSKGLIPLSS	1399

RESULT 10

US-09-388-221-6
; Sequence 6, Application US/09388221A
; Publication No. US20020192643A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; TITLE OF INVENTION: No. US20020192643A1el Card Proteins Involved in Cell Death Regul
; FILE REFERENCE: P-LJ 3650
; CURRENT APPLICATION NUMBER: US/09/388,221A
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1443
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-388-221-6

Query Match 97.3%; Score 7332; DB 10; Length 1443;
Best Local Similarity 95.0%; Pred. No. 0;
Matches 1399; Conservative 0; Mismatches 0; Indels 74; Gaps 2;
Qy 1 MAGGAWRLACYLEFLFKELKEFOLLANKAHSSSSGETTAQPKTSGMEVASLYVAQ 60
Db 1 MAGGAWRLACYLEFLFKELKEFOLLANKAHSSSSGETTAQPKTSGMEVASLYVAQ 60

Qy	61	YGEQAWDLALHTWEOMGLRSICAQOEGAGHSPPSPSEPHLGSPSOPTSTAVLMPW	120
Db	61	YGEQAWDLALHTWEOMGLRSICAQOEGAGHSPPSPSEPHLGSPSOPTSTAVLMPW	120
Qy	121	IHELPAQCTQCSERRVLRQLPDTSGRRWRREISASLLYQALPSSPDHSPSPQESPNAPTST	180
Db	121	IHELPAQCTQCSERRVLRQLPDTSGRRWRREISASLLYQALPSSPDHSPSPQESPNAPTST	180
Qy	181	AVLSGWSGPPQPSLAPREOEAPGTOWPLDETSGIYYTIREREREKSEKGRPPMAAVGT	240
Db	181	AVLSGWSGPPQPSLAPREOEAPGTOWPLDETSGIYYTIREREREKSEKGRPPMAAVGT	240
Qy	241	PPQAHTSLQPHHHHPWSPVRESLSTWPKNEDFNQKFTOLLQRPHPRSODPLVKRWS	300
Db	241	PPQAHTSLQPHHHHPWSPVRESLSTWPKNEDFNQKFTOLLQRPHPRSODPLVKRWS	300
Qy	301	PDYVEENRGHLIEIRDLFGPLDQEPRIVLQGAAGIGKSTLARQVKEAWGRGOLYGRD	360
Db	301	PDYVEENRGHLIEIRDLFGPLDQEPRIVLQGAAGIGKSTLARQVKEAWGRGOLYGRD	360
Qy	361	FQHVFFYFSCRELAQSKVSVLAELIGKDGATATPAPIRQILSRPERLLFLDGVDEPGWVLQ	420
Db	361	FQHVFFYFSCRELAQSKVSVLAELIGKDGATATPAPIRQILSRPERLLFLDGVDEPGWVLQ	420
Qy	421	EPSSCLCHWSQPADALLGSLGKTLTPENASFLITARTTALQNLIPSLEQARVVEVLG	480
Db	421	EPSSCLCHWSQPADALLGSLGKTLTPENASFLITARTTALQNLIPSLEQARVVEVLG	480
Qy	481	FSESRKEVYFYFDERQAIARFRLVKSNNELWALCLVPVWSWLACTCLMOMKREKL	540
Db	481	FSESRKEVYFYFDERQAIARFRLVKSNNELWALCLVPVWSWLACTCLMOMKREKL	540
Qy	541	TLTSTTTTTCLHYLAQALQAPLQPLQDLCLSLAAEGIWQKTLFSPDDLKKGIDGAI	600
Db	541	TLTSTTTTTCLHYLAQALQAPLQPLQDLCLSLAAEGIWQKTLFSPDDLKKGIDGAI	600
Qy	601	ISTFLKMGILQHPILPSYFIHLFCQFFFAAMSVVLEDEKGRGHKSNCIIDLEKTLEY	660
Db	601	ISTFLKMGILQHPILPSYFIHLFCQFFFAAMSVVLEDEKGRGHKSNCIIDLEKTLEY	660
Qy	661	GIHGLFGASTTRFLGLLSDGEREMENIFHCLRSQGRNLMQVPSLQLLQPHLSLH	720
Db	661	GIHGLFGASTTRFLGLLSDGEREMENIFHCLRSQGRNLMQVPSLQLLQPHLSLH	720
Qy	721	CLYETRNKFTLTQVMAHFEEMGCVETDMLLVCTFCIKFSRHVKQLQLIEGRQHRSTWS	780
Db	721	CLYETRNKFTLTQVMAHFEEMGCVETDMLLVCTFCIKFSRHVKQLQLIEGRQHRSTWS	780
Qy	781	PTMVLFWRVPTVDAYWQILFSLVKVTRNLKELDLGNSLSHSAVKSCKTLRRPRCLLE	840
Db	781	PTMVLFWRVPTVDAYWQILFSLVKVTRNLKELDLGNSLSHSAVKSCKTLRRPRCLLE	840
Qy	841	TURLAGCGLTABDCCKDLAFGLRANOTLTLDLSFNVLTDAGAKHLCORLQPSCKLQRLQ	900
Db	841	TURLAGCGLTABDCCKDLAFGLRANOTLTLDLSFNVLTDAGAKHLCORLQPSCKLQRLQ	900
Qy	901	LVSCLGTSDDCCDLASVLSASPSLKDLDQNNLDVGVRLLCCEGLRHLPACKLIRLGLDQ	960
Db	901	LVSCLGTSDDCCDLASVLSASPSLKDLDQNNLDVGVRLLCCEGLRHLPACKLIRLGLDQ	960
Qy	961	TTLSDMRQELRALEQEKPOLLIIFSRKFPVMTPTTEGLDTGEMSNSTSLKQRLGSERA	1020
Db	958	-----KFSVMTPTTEGLDTGEMSNSTSLKQRLGSERA	990
Qy	1021	ASHVAQANLKLDVSKIPPIAIEBESSPEVVPVELLCVPSASQGDHLTKPLGTTDDDFW	1080
Db	991	ASHVAQANLKLDVSKIPPIAIEBESSPEVVPVELLCVPSASQGDHLTKPLGTTDDDFW	1050
Qy	1081	GPTGPVATEVVDKKNLYRVHPFVAGSVRWPNVTGLCFVNRVAVTVEIEFCVWDQFLGEIN	1140
Db	1051	GPTGPVATEVVDKKNLYRVHPFVAGSVRWPNVTGLCFVNRVAVTVEIEFCVWDQFLGEIN	1110
Qy	1141	POHSMWVAGPLLDIKAEPGAVEAVHLPHFVALQGGHVDTSLFQMAHFEKGMLEKPARV	1200

Db	1111	PQHSNMVAGPLDDIKAEPCAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPARV	1170
Qy	1201	ELHHIVLENPSPLGVLKMIHNAIRFPVTSVLLYHRVHPPEVTFHLYLIPSDCSIR	1260
Db	1171	ELHHIVLENPSPLGVLKMIHNAIRFPVTSVLLYHRVHPPEVTFHLYLIPSDCSIR	1230
Qy	1261	-----KELELCYRSPGSDQLF	1276
Db	1231	KAIDDLKMPQFVRIHKPPPLTPLYNGCRYTVSGSGSMLEILPKELELCYRSPGSDQLF	1290
Qy	1277	SEFYVGHLSGIRLQVKDKDETLVWEALVKPGDLMPATTLIPPARIAVPSPLDAPQLLH	1336
Db	1291	SEFYVGHLSGIRLQVKDKDETLVWEALVKPGDLMPATTLIPPARIAVPSPLDAPQLLH	1350
Qy	1337	FVDQYREQLIARVTSVEVVDLKHGOVLQSOEQYERVAENTRPSQMRKLFSLSQSWDRKC	1396
Db	1351	FVDQYREQLIARVTSVEVVDLKHGOVLQSOEQYERVAENTRPSQMRKLFSLSQSWDRKC	1410
Qy	1397	KDGLYQALKETHPHLIMELWEKSGKGLPLSS	1429
Db	1411	KDGLYQALKETHPHLIMELWEKSGKGLPLSS	1443

RESULT 11

US-09-388-221-10

Sequence 10, Application US/09388221A

Publication No. US20020192643A1

GENERAL INFORMATION:

APPLICANT: Reed, John C.

TITLE OF INVENTION: No. US20020192643A1el Card Proteins Involved in Cell Death Regul

FILE REFERENCE: P-LT 3650

CURRENT APPLICATION NUMBER: US/09/388,221A

CURRENT FILING DATE: 1999-09-01

NUMBER OF SEQ ID NOS: 18

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 10

LENGTH: 1454

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic

US-09-388-221-10

Query Match 84.2%; Score 6344; DB 10; Length 1454;

Best Local Similarity 84.2%; Pred. No. 0;

Matches 1229; Conservative 53; Mismatches 122; Indels 56; Gaps 7;

Qy	1	MAGGAWRLACYLEFLKKEELKEFOLLANKAHSRSSSGETPAQPKTSGMEVASYLVAQ	60
Db	1	MAGGAWRLACYLEFLKKEELKEFOLLANKAHSRSSSGETPAQPKTSGMEVASYLVAQ	60
Qy	61	YGEQAWDLALHTWEQMLRSLSQAQAGAGHSPSPYPSPHGLSPSQPTSTAVLMPW	120
Db	61	YGEQAWDLALHTWEQMLRSLSQAQAGAGHSPSPYPSPHGLSPSQPTSTAVLMPW	120
Qy	121	IHELPAQTQGGERRVRLQDPTSGRRWREISASLLYQALPSDPHSPSQSNAPTST	180
Db	121	IHELPAQTQGGERRVRLQDPTSGRRWREISASLLYQALPSDPHSPSQSNAPTST	180
Qy	181	AVLWSGSPPPQSLAPREQEAPQOWPLDETSGIYYTEIREREREKSEKRPMPAAVGT	240
Db	181	AVLWSGSPPPQSLAPREQEAPQOWPLDETSGIYYTEIREREREKSEKRPMPAAVGT	240
Qy	241	PPQAHTSLQPHHPHPSPVRESLCSLTPWPKNEDFNQKFTQLLLQRPHPRSQDPLVRSW	300
Db	241	PPQAHTSLQPHHPHPSPVRESLCSLTPWPKNEDFNQKFTQLLLQRPHPRSQDPLVRSW	300
Qy	301	PDYVEENRGLHIEIRDLFGPLDTPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGR	360
Db	301	PDYVEENRGLHIEIRDLFGPLDTPRIVILQGAAGIGKSTLARQVKEAWGRGQLYGR	360
Qy	361	FQHVFFYFSCRELAQSKVSVSLAEILIGKDGATATPAPIRQILSRPERLLFILDGVDPCGWVLQ	420

Db	361	FQHVFFYFSCRELAQSKVSVSLAEILIGKDGATATPAPIRQILSRPERLLFILDGVDPCGWVLQ	420
Qy	421	EPSSSCLCHWSQPADALLGSLGKTKTILPEASFLITARTTALQNLIFSLQEARVEVLG	480
Db	421	EPSSSCLCHWSQPADALLGSLGKTKTILPEASFLITARTTALQNLIFSLQEARVEVLG	480
Qy	481	PSESSRKEVYFYFDERQAIAPRFLVKSNEKLMALCLVPMVSWLACTCLMOQMRKEKL	540
Db	481	PSESSRKEVYFYFDERQAIAPRFLVKSNEKLMALCLVPMVSWLACTCLMOQMRKEKL	540
Qy	541	TLTSTTTTTTCLCHYLAQALQAQPLGQPLRDLCSLAAEGIWQKTKTILFSPDDLRKHGLDGA	600
Db	541	TLTSTTTTTTCLCHYLAQALQAQPLGQPLRDLCSLAAEGIWQKTKTILFSPDDLRKHGLDGA	600
Qy	601	ISTFLKMGILQEHPIPLSYFTHLCFQBFPAAMSIVLEDEKGRGHKSNCCIIDEKTLAY	660
Db	601	ISTFLKMGILQEHPIPLSYFTHLCFQBFPAAMSIVLEDEKGRGHKSNCCIIDEKTLAY	660
Qy	661	GIHGLFGASTTRFLGLLSDGEGEREMENIFHCRLSQGRNLMQWVPSLQLLQPHSLES	720
Db	661	GIHGLFGASTTRFLGLLSDGEGEREMENIFHCRLSQGRNLMQWVPSLQLLQPHSLES	720
Qy	721	CLYETRNTKTLTQVNAHFEEMCMVETDMLLVCTFCIKFSRHVKKLQLEGRQHRSTWS	780
Db	721	CLYETRNTKTLTQVNAHFEEMCMVETDMLLVCTFCIKFSRHVKKLQLEGRQHRSTWS	780
Qy	781	PTMVVLFPMVPTDAYWQILFSLVLTNRNLKELDLSGNSLSHSAVKSICKTLRRPRCLLE	840
Db	781	PTMVVLFPMVPTDAYWQILFSLVLTNRNLKELDLSGNSLSHSAVKSICKTLRRPRCLLE	840
Qy	841	TURLAGCGLTAECDCKDLAFGRANOTLTDELDSFNLVTDAGAKHLQRLQPSCKLQRLQ	900
Db	841	TURLAGCGLTAECDCKDLAFGRANOTLTDELDSFNLVTDAGAKHLQRLQPSCKLQRLQ	900
Qy	901	LVSCGLTSDCCODLASVLSASPSELKELDLQNNLDDVGVRLICEGRHPACKLIRGLDQ	960
Db	901	LVSCGLTSDCCODLASVLSASPSELKELDLQNNLDDVGVRLICEGRHPACKLIRGLDQ	960
Qy	961	TTLSDMRQELRALRQEKPOLLIIFSRKPSVMTPTGLDGTGEMSNSTSLKQRLGSERA	1020
Db	961	TTLSDMRQELRALRQEKPOLLIIFSRKPSVMTPTGLDGTGEMSNSTSLKQRLGSERA	1020
Qy	1021	ASHVAQANILLDVSKIPPIAEIAEESSEVPVVELLCVPSASQGLHTKPLGTDGDFW	1080
Db	1021	ASHVAQANILLDVSKIPPIAEIAEESSEVPVVELLCVPSASQGLHTKPLGTDGDFW	1080
Qy	1081	GPTGPVATEVVDKKNLYRHPVFPVAGSYRWPNTGLCFVMREAVTVEIEFCVMDQFLG-EI	1139
Db	1081	GPEGNVDELIDKSTNRYSVMPPTAGWYLSATGLGLVRDEVTVTIAFGWSQHLALDL	1140
Qy	1140	NPQHSNMVAGPLDDIKABP-GAVEAVHLPHFVALQGGHVDTSLFQMAHFKEEGMLLEKPA	1198
Db	1141	QHHEQMLVGGPLFDVTABPEEPAVAIEHLPHFISLQ-GEVDYSWFLVAHFKQEGMVLHPA	1199
Qy	1199	RVELHHIVLENPSPLGVLKMIHNAIRFPVTSVLLYHRVHPPEVTFHLYLIPSDCS	1258
Db	1200	RVEPPYAVLESFSPSLMGILLRIASGTRLSPTTNTLIYYHPHEDIKFHLIYLPDQAL	1259
Qy	1259	IR-----KELELCYRSPGSDQLF	1276
Db	1260	LTKAIDDEDRPHGVRLQTSPPMEPLNFGSSYIVSNSANLKVMPKELKLSYRSPGEIQHF	1319
Qy	1277	SEFYVGHLSGIRLQVKDKDETLVWEALVKPGDLMPATTLIPPARIAVPSPLDAPQLLH	1336
Db	1320	SKFYAGQMKETQLEITEKRHGTLMVDTKVPVLDLQVLAASAPP-----PFGAA---	1369
Qy	1337	FVDQYREQLIARVTSVEVVDLKH-GOVLQSOEQYERVAENTRPSQMRKLFSLSQSWDRK	1395
Db	1370	FVKNHRQLQARMGDLKGVLDLQDNVEITENEKELVEQEKTRQSKNEALLSWBKGD	1429
Qy	1396	CKDGLYQALKETHPHLIMEL	1415

Db 1430 ALDVLFRSISERDPYLVSYL 1449

RESULT 12

US-09-388-221-12

; Sequence 12, Application US/09388221A

; Publication No. US20020192643A1

; GENERAL INFORMATION:

; APPLICANT: Reed, John C.

; TITLE OF INVENTION: P-1J 3650

; FILE REFERENCE: US/09/388, 221A

; CURRENT APPLICATION NUMBER: US/09-09-01

; CURRENT FILING DATE: 1999-09-01

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 12

; LENGTH: 1424

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Synthetic

US-09-388-221-12

Query Match 81.9%; Score 6174; DB 10; Length 1424;

Best Local Similarity 82.1%; Pred. No. 0;

Matches 1199; Conservative 53; Mismatches 122; Indels 86; Gaps 8;

QY 1 MAGGAWRLACYLEFLKKEELKEFOLLANKAHSRSSSGETPAQPEKTSQMEVASVYLAQ 60

Db 1 MAGGAWRLACYLEFLKKEELKEFOLLANKAHSRSSSGETPAQPEKTSQMEVASVYLAQ 60

QY 61 YGEQAWDLALHTWQMLRSLSQAQEGAGHSPGFPYSPPHLGSQSQTSTAVLMPW 120

Db 61 YGEQAWDLALHTWQMLRSLSQAQEGAGHSPGFPYSPPHLGSQSQTSTAVLMPW 120

QY 121 IHELPAQTQGSERRVRLPDTSGRRWREISASLLYQALPSPDHESPSQSNAPTST 180

Db 121 IHELPAQTQGSERRVRLPDTSGRRWREISASLLYQALPSPDHESPSQSNAPTST 180

QY 181 AVLGWSGSPQSLAPREQEAGCTQPLDETSGIYYTEIREREREKSEKGRPPMAAVGT 240

Db 181 AVLGWSGSPQSLAPREQEAGCTQPLDETSGIYYTEIREREREKSEKGRPPMAAVGT 240

QY 241 PQAHTSLQPHHPHPPSPVRESLCSWPKWNEFQKFTQLLLQRPSPRSDPLVKRSW 300

Db 241 PQAHTSLQPHHPHPPSPVRESLCSWPKWNEFQKFTQLLLQRPSPRSDPLVKRSW 300

QY 301 PDYVENRGLHLEIRDLFGGLDQTPRIVILQGAAGICKSTLARQVKEAWGRGQLYGDR 360

Db 301 PDYVENRGLHLEIRDLFGGLDQTPRIVILQGAAGICKSTLARQVKEAWGRGQLYGDR 360

QY 361 FOHVYFSCRELAQSKVWSLAELIGKDGATAPIRQILSRPERLLFILDGDEPGWVLO 420

Db 361 FOHVYFSCRELAQSKVWSLAELIGKDGATAPIRQILSRPERLLFILDGDEPGWVLO 420

QY 421 EPSSELCHWSQPQADALLGSLGHTILPEASFLITARTTALQNLIPSEQARWVEVLG 480

Db 421 EPSSELCHWSQPQADALLGSLGHTILPEASFLITARTTALQNLIPSEQARWVEVLG 480

QY 481 FSESSRKEYFYRTDERQAIRAFRLVKSNNKELWALCLVPWVSWLACTCLMQMKRKEKL 540

Db 481 FSESSRKEYFYRTDERQAIRAFRLVKSNNKELWALCLVPWVSWLACTCLMQMKRKEKL 540

QY 541 TLTSKTTTLLCHLYLAQALQAPGLQRLDLSLAAGIWQKTLFSPDDLRKHGLDGA 600

Db 541 TLTSKTTTLLCHLYLAQALQAPGLQRLDLSLAAGIWQKTLFSPDDLRKHGLDGA 600

QY 601 ISTFLKMGILQEHPIPLSYFTHLCQFEFFAAMSVYLEDEKGRGHSNCIIDEKLTLEY 660

Db 601 ISTFLKMGILQEHPIPLSYFTHLCQFEFFAAMSVYLEDEKGRGHSNCIIDEKLTLEY 660

QY 661 GIHGLFGASTTRFLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLOPHSLES 720

Db 661 GIHGLFGASTTRFLGLLSDEGEREMENIFHCRLSQGRNLMQWVPSLQLLLOPHSLES 720

QY 721 CLYETRNKTFLTQVMAHPEEMCMCVETDMELLVCTFCIKFSRHVKKLQLEIGRQHRSTWS 780

Db 721 CLYETRNKTFLTQVMAHPEEMCMCVETDMELLVCTFCIKFSRHVKKLQLEIGRQHRSTWS 780

QY 781 PTMVVLFWRVPTDAYWQILFSLVKVTRNLKELDLSGNSLSHSAVSKLCKTLRRPRCLLE 840

Db 781 PTMVVLFWRVPTDAYWQILFSLVKVTRNLKELDLSGNSLSHSAVSKLCKTLRRPRCLLE 840

QY 841 TLRLAGCGLTAEDCKDLAFGRANOTLTLELDLSFNVLTDAGAKHLQRLRQPSCKLQRLQ 900

Db 841 TLRLAGCGLTAEDCKDLAFGRANOTLTLELDLSFNVLTDAGAKHLQRLRQPSCKLQRLQ 900

QY 901 LVSCGLTSDCCQDLASVLSASPSELKELDLQNNLDDVGVRLLCCEGLRHPACKLIRGLDQ 960

Db 901 LVSCGLTSDCCQDLASVLSASPSELKELDLQNNLDDVGVRLLCCEGLRHPACKLIRGLDQ 960

QY 961 TTLSDEMQRALAEQEKPOLLIIFSRKPSVMTPTTEGLDTGEMSNSTSSLRQRQSGSRA 1020

Db 958 -----KPSVMTPTTEGLDTGEMSNSTSSLRQRQSGSRA 990

QY 1021 ASHVAQANLKLDSKIPPIAEIAEBSSEPVVVELLCVPSASQGDHLTKPLGTDGDFW 1080

Db 991 ASHVAQANLKLDSKIPPIAEIAEBSSEPVVVELLCVPSASQGDHLTKPLGTDGDFW 1050

QY 1081 GPTGVATEVVDKKNLYRVHFPVAGSVRWNTGLCFVMREAVTVEIEFCVMDQFLG-EI 1139

Db 1051 GPEGNVDELIDKSTNRYSVWNPFTAGWYLSATGLGLVRDEVVTIAFGSWSQHLALDL 1110

QY 1140 NPQHSMMVAGPLDIDKAEF-GAVEAVHLPHEFVALQGGHVDTSLFQMAHFKEBGMLEKPA 1198

Db 1111 QHHEQWLVGGPLDFVTAEPPEAEVAEIHLPHEFISLQ-GEVDVSWFLVAHFKNMGVLEHPA 1169

QY 1199 RVELHHVLENPSPGLVLLKMHNLARFIPVTSVLLYHRVPEETVHLYLIPSDCS 1258

Db 1170 RVEPFYVLESFPSFLMGILLIRIASGTRLSIPTNTLIYYHPHEDIKFHLVLPESDAL 1229

QY 1259 IR-----KELELCYRSPGEDQLF 1276

Db 1230 LTKAIDDEEDRPHGVRLQTSPPMEPLNFGSSYIVSNSANLKVMPKELKLSVRSPEIQHF 1289

QY 1277 SEFYVHGLSGIRLQVKKDDTLVWEALVKPGDLMPTTILPPARIAVPSPLDAPQLLH 1336

Db 1290 SKFYAGOMKEPQLEITEKRHGTGLVMDTEVPKVDQLVAASAPP-----PFSGAA--- 1339

QY 1337 FVDQVREQLIARVTSVEVVLKHLH-GQVLSQEQYERVLAEINTRPSQMRKLPFSLQSQSWDRK 1395

Db 1340 FVKENHRQLQARMGDLKGLVDLQDNVELTENEKELVEQEKTRQSKNEALLSMVEKKGD 1399

QY 1396 CKDGLYQALKETHPHLIMEL 1415

Db 1400 ALDVLFRSISERDPYLVSYL 1419

RESULT 13

US-10-407-866-92

; Sequence 92, Application US/10407866

; Publication No. US20040002593A1

; GENERAL INFORMATION:

; APPLICANT: Reed, John C.

; APPLICANT: Godzik, Adam

; TITLE OF INVENTION: PAD Domain-Containing Polypeptides, and Methods of Use

; FILE REFERENCE: 66654-10 (UJ 5755)

; CURRENT APPLICATION NUMBER: US/10/407,866

; CURRENT FILING DATE: 2003-04-04

; PRIOR APPLICATION NUMBER: US 60/370,538

; PRIOR FILING DATE: 2002-04-04

; NUMBER OF SEQ ID NOS: 129

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 92

; LENGTH: 764

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-407-866-92

Query Match      54.0%; Score 4072; DB 12; Length 764;
Best Local Similarity 100.0%; Pred. No. 1.3e-314;
Matches 764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGGAWRLACYLEFLKKEBELKEFQLLANKAHSRSSSGETPAQPKTSGMEVASYLVAQ 60
Db 1 MAGGAWRLACYLEFLKKEBELKEFQLLANKAHSRSSSGETPAQPKTSGMEVASYLVAQ 60
QY 61 YGEQAWDLALHTWEQMLSLCAQAEAGAGHSPSPYPSPSEPHLGSPSQPTSTAVLMPW 120
Db 61 YGEQAWDLALHTWEQMLSLCAQAEAGAGHSPSPYPSPSEPHLGSPSQPTSTAVLMPW 120
QY 121 IHELPAGCTQGSERRVLRQLPDTSGRRWEISASLLYQALPSPDHESQSPNAPTST 180
Db 121 IHELPAGCTQGSERRVLRQLPDTSGRRWEISASLLYQALPSPDHESQSPNAPTST 180
QY 181 AVILGWSGPPQSLAPREOEAPCTOWPLDETSGIYYTEIREREREKSGRPPMAAVGT 240
Db 181 AVILGWSGPPQSLAPREOEAPCTOWPLDETSGIYYTEIREREREKSGRPPMAAVGT 240
QY 241 PQAHTSLQPHHPHWPSPVRESLCSTWPKNEDFNQFTQLLLQRPSPRQDPLVKRSM 300
Db 241 PQAHTSLQPHHPHWPSPVRESLCSTWPKNEDFNQFTQLLLQRPSPRQDPLVKRSM 300
QY 301 PDYVENRHLIEIRDLFGGLDTPQPRIVILQGAAGIKSTLARQVKWARGQLYGR 360
Db 301 PDYVENRHLIEIRDLFGGLDTPQPRIVILQGAAGIKSTLARQVKWARGQLYGR 360
QY 361 FOHFVYFSCRELAQSKVSLAEILGKGTATPAPRQILSRPERLLFILDGVDPEGVLIQ 420
Db 361 FOHFVYFSCRELAQSKVSLAEILGKGTATPAPRQILSRPERLLFILDGVDPEGVLIQ 420
QY 421 EPSSELCLHWSQPQADALGSLGKTIILPEASFLITARTTALQNLIPSEQARWVEVLG 480
Db 421 EPSSELCLHWSQPQADALGSLGKTIILPEASFLITARTTALQNLIPSEQARWVEVLG 480
QY 481 FSESSKEYFYRTDERQAIRAFRLVKNKELWALCLVPWVSWLACTCLMQMKRKEKL 540
Db 481 FSESSKEYFYRTDERQAIRAFRLVKNKELWALCLVPWVSWLACTCLMQMKRKEKL 540
QY 541 TLTSKTTTTLCHLYLAQAQPLGQPLRDCLSLAEGIWQKTLFSPDLRKHGLDGA 600
Db 541 TLTSKTTTTLCHLYLAQAQPLGQPLRDCLSLAEGIWQKTLFSPDLRKHGLDGA 600
QY 601 ISTFLKMGILQEHPIPLSYSFHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660
Db 601 ISTFLKMGILQEHPIPLSYSFHLCFQEFFAAMSYVLEDEKGRGKHSNCIIDLEKTLEAY 660
QY 661 GIHGLFGASTTRFLGLLSDGEREMENIFHCRLSQGRNLMQWVPSLQLLLPHSLESIH 720
Db 661 GIHGLFGASTTRFLGLLSDGEREMENIFHCRLSQGRNLMQWVPSLQLLLPHSLESIH 720
QY 721 CLYETRNKTLFQVMAHPEMGVCVETDMLLVCTFCIKFSRHV 764
Db 721 CLYETRNKTLFQVMAHPEMGVCVETDMLLVCTFCIKFSRHV 764

RESULT 14
US-09-895-298-139
; Sequence 139, Application US/09895298
; Publication No. US20030078405A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 47 Human Secreted Proteins
; FILE REFERENCE: P2035P1
; CURRENT APPLICATION NUMBER: US/09/895,298
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: 09/591.16
; PRIOR FILING DATE: 2000-06-09

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-407-866-68

Query Match      28.8%; Score 2171.5; DB 11; Length 442;
Best Local Similarity 97.4%; Pred. No. 8.4e-164;
Matches 414; Conservative 4; Mismatches 4; Indels 3; Gaps 1;

QY 997 GLDTGEMSNSTSSLRQRLGSEAAASHVAQANLKLDDVSKIFPIAETAEESPEVVPVEL 1056
Db 1 GLDTGEMSNSTSSLRQRLGSEAAASHVAQANLKLDDVSKIFPIAETAEESPEVVPVEL 60
QY 1057 LCVPSPASQGLDHTKPLGTDDDFWGPPTGPVATEVVDKKNLYRVHFPVAGSYRWPNTGLC 1116
Db 61 LCMPSPASQGLDHTKPLGTDDDFWGPPTGPVATEVVDKKNLYRVHFPVAGSYRWPNTGLC 120
QY 1117 FVMREAVTVEIEFCVWDQFLGEINPQHSMMVAGPLDDIKAEPPGAVEAVHLPHFVALQGGH 1176
Db 121 FVMREAVTVEIEFCVWDQFLGEINPQHSMMVAGPLDDIKAEPPGAVEAVHLPHFVALQGGH 180
QY 1177 VDTSLFQMAHFKEEGMLLEKPARVELHHIVLENFSPFSLGVLKMIHNAIRFIPVTSVVL 1236
Db 181 VDTSLFQVAFHKEEGMLLEKPARVELHHIVLENFSPFSLGVLKMIHNAIRFIPVTSVVL 240
QY 1237 LYHRVHPEVTFHLYLIPSDCSIRKELELCYRSGEDQDFSEFYVGHGSLGIRLOVKDKK 1296
Db 241 LYHRVHPEVTFHLYLIPSDCSIRKELELCYRSGEDQDFSEFYVGHGSLGIRLOVKDKK 300
QY 1297 DETLVWEALVPGDLMPTATLIPARIAPSPDAPQLLHFVDQYREQLIARVTSVEVVL 1356
Db 301 DETLVWEALVPGDLMPTATLIPARIAPSPDAPQLLHFVDQYREQLIARVTSVEVVL 360
QY 1357 DKLHGQVLSQBYERVAENRTPSQMRKLFSLSQSWDRKCKDGLYQALKETHPHLMELW 1416
Db 361 DKLHGQVLSQBYERVAENRTPSQMRKLFSLSQSWDRKCKDGLYQALKETHPH 417
QY 1417 EKGSK 1421
Db 418 NSGRR 422

RESULT 15
US-10-407-866-68
; Sequence 68, Application US/10407866
; Publication No. US20040002593A1
; GENERAL INFORMATION:
; APPLICANT: Reed, John C.
; APPLICANT: Godzik, Adam
; TITLE OF INVENTION: PAAD Domain-Containing Polypeptides,
; TITLE OF INVENTION: Encoding Nucleic Acids, and Methods of Use
; FILE REFERENCE: 66654-10 (LJ 5755)
; CURRENT APPLICATION NUMBER: US/10/407,866
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US 60/370,538
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68
; LENGTH: 1027
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-407-866-68

```


Query Match 18.4%; Score 1382.5; DB 12; Length 1027;
Best Local Similarity 34.4%; Pred. No. 1.9e-100;
Matches 371; Conservative 143; Mismatches 373; Indels 193; Gaps 24;

Qy	8	RLACYLEFLKKEFKFQFLLANKAHSSSSGETP-AQBEKTSGMVEASVLAQYCEQRA	66
Db	12	RUSTYLEELEAEVLFKFKLYLGTG--TELGEKIPWGSMEKAGPLEMAQLLITHEGPEEA	69
Qy	67	WDLALHTWQMGILRSICAQAGAGHSPFPYSPSEPHLGSPSQ-PTSTAVLMPWIHELP	125
Db	70	WRLALSTFERINPKDLWERCQREDLVDPGPGSS--LGNQSTCLLEVLVTP--RKDP	125
Qy	126	AGCTQGSERRVLQPLDPTSGRRWREISALLVOALPSSDHSPSQESPNAPTSTAVLGS	185
Db	126	QETYRDYVRKFRLMEDRNARLGEVCNLSHRVTRLLLVKEHSNPMQVQ-----	173
Qy	186	WGSPPQPSLAPQEQAPGTQWPLDETSGIYYTEIREREREKSEKGPWWAAVVGTPPOAH	245
Db	174	-----	173
Qy	246	TSIQPHHPWEPSPVRESLCTWPKWNEFNQKFTQLLLQRPHPRSQDPLVKRSMYPDVE	305
Db	174	-----QQLLDTGRGHARTV-----	187
Qy	306	ENRGHL---IEIRDLPGLDQOE-PRIVLQGAAGIGKSTLARQVKEAWGRQLYGDRF	361
Db	188	---GHQASPIKIBTLPEPDEERPEPRTVMQGAAGIGKSMLAHKVMDWADGKLFQGRF	244
Qy	362	QHVYFSCRELAQSKV-VSLAELIGKDGATAPAPIQLSRPERLLFILDGVDEPGWVLO	420
Db	245	DYLFYINCREMNSATECSMDLIIFSCEPSPAPLQELIRVPERLLFIIDGDELKPSFH	304
Qy	421	EPSELCLHWSQOPDADLGLSGTLILPEASFLITARTALONLIPSLQARWVEVLG	480
Db	305	DPQCPWCLWEEXRPTLELLNSLRKLLPELSELLITRTALEKLRHLEHPRHVEILG	364
Qy	481	FSSESRKEYFYRYFTDERQAIRAFRLVKNGKELWALCLVPVWSWLACTCLMQMKRKEKL	540
Db	365	FSEAERKEYFYKVFHNAEQAGQVFNVRDNEPLFTFCVPLVCWVCTCLQQLEGGGLL	424
Qy	541	TLTSKTTTTLCHYLALQALQAPLGPOL-----ROLCSLAEGIWQKXKTLFSPDDLKH	594
Db	425	RQTSRTTAVYMLYLLSLMQPKPGAPRLQPPNQRLGCLSLAADGLWNQKILFEEQDLKH	484
Qy	595	GLDGAIISTFLKMGILQEH-PIPLSYSFTHLCFQEPFAAMSYYL-EDEKGRGKHSNCIID	652
Db	485	GLDGEDVSAPLANNIIFQKINCERYYSFTHLSFQEFPAAMYIILDEGEGAGDPQ-----D	540
Qy	653	LEKTLFAYGI-HGLFGASTTRFLLGLLSDEGEREMENIFHCRLSQ--GRNLMQWVPSLQ-	708
Db	541	VTRLLTEYAFSPERSFLATLSRFLFGLLNEETRSHLEKSLCWKVSPIHKMDLLQWISKAQ	600
Qy	709	---LLLOPHSLSLHCLYETRNKTLTQVVAHFEEMGMC-VETDMELLVCTFCIKFSRHV	764
Db	601	SDGSTLQQGSLEFSCLYEIQEEFIQALSHFQVIWVSNIASMEHVMVSSFCLKCRSA	660
Qy	765	KILQLI-----EGRQHSTWSPWVLF--WVPVTDAYWQILFVSLKVTRNLKELD	814
Db	661	QVLHLGYATYSADGEDRARCAGAHLLVQLRPRTVLLDAYSEHLAAALCTNPNLIELS	720
Qy	815	LSGNSLSHSAVKSLCKTLRRPRCLLETRLLAGCGLTAEDCKDLAFGLRANQTLTELDLSF	874
Db	721	LYRNALGSRGVKLLCQGLRHPNCKLQNLRLKRCRISSACEDLSAALIANKNLTRMDLSG	780
Qy	875	NVLTDAKAKHLQRLRQPSCKLQRLQVSCGLTSDCCQDLASVLSASPSLKELDLOONNL	934
Db	781	NGVGFPGMMLLCEGLRHPQRLQMIQLRKQLESQACQEMASVLTGNPHLVELDLTGNAL	840
Qy	935	DVGVRLLCGLRHPACKLIRLGL-----DQTLSDEMROELRALEQEKPOLLIIF	984
Db	841	EDLGLRLLCQGLRHPVCRRLTLWLKICRLTAACADELASTLSYNQSLRELDLSNEL---	897
Qy	985	SRKPSVMTPEGLDTGEMNSNSTSLKRQRLGSEAAASHV-----AQANLKLLDVS	1035

Db 898 --GDLGVILLCEGL-----RHPTCKLQTLRLGICRLGSAACEGLSVLQANHNRELDLS 950
Search completed: January 29, 2004, 13:54:37
Job time : 252.019 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2004, 02:57:41 ; Search time 2662.52 Seconds
(without alignments)
11370.118 Million cell updates/sec

Title: US-09-996-617-7
Perfect score: 740
Sequence: 1 cgcgtccggctgcagcgggg.....atacgaagccagcttgaa 740

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sta.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sta.*

28: em_un.*

29: em_vl.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_man.*

37: em_htg_vrt.*

38: em_sv.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	740	100.0	740	6	AR256267	AR256267 Sequence
2	740	100.0	740	6	AX082246	AX082246 Sequence
3	740	100.0	740	6	AX536208	AX536208 Sequence
4	740	100.0	740	6	AX536223	AX536223 Sequence
5	740	100.0	740	9	AF384665	AF384665 Homo sapi
6	734	99.2	770	6	AX118619	AX118619 Sequence
7	734	99.2	770	9	AF184073	AF184073 Homo sapi
8	734	99.2	782	6	AX459863	AX459863 Sequence
9	734	99.2	782	9	AB023416	AB023416 Homo sapi
10	733	99.1	779	6	AX017270	AX017270 Sequence
11	733	99.1	779	6	AX524974	AX524974 Sequence
12	733	99.1	779	6	BD134441	BD134441 Human nuc
13	730.8	98.8	785	6	BD057255	BD057255 Apoptosis
14	713	96.4	740	9	AK000211	AK000211 Homo sapi
15	687	92.8	712	9	AF310103	AF310103 Homo sapi
16	680	91.9	772	9	BC013569	BC013569 Homo sapi
17	610	82.4	713	6	AX118641	AX118641 Sequence
18	610	82.4	713	9	AF255794	AF255794 Homo sapi
19	585	79.1	585	6	AR256268	AR256268 Sequence
20	585	79.1	585	6	AX082248	AX082248 Sequence
21	585	79.1	585	6	AX536210	AX536210 Sequence
22	448	60.5	450	12	BT007571	BT007571 Synthetic
23	437.6	59.1	768	9	BC004470	BC004470 Homo sapi
24	416.8	56.3	588	4	AB050006	AB050006 Bos tauru
25	378	51.1	405	6	AX118643	AX118643 Sequence
26	360	48.6	2821	6	AX118618	AX118618 Sequence
27	360	48.6	2821	9	AF184072	AF184072 Homo sapi
28	360	48.6	127769	9	AC009088	AC009088 Homo sapi
29	356	48.1	356	6	AX118626	AX118626 Sequence
30	336	45.4	732	10	BC008252	BC008252 Mus muscu
31	336	45.4	770	10	AB032249	AB032249 Mus muscu
32	336	45.4	777	6	AR256273	AR256273 Sequence
33	336	45.4	777	6	AX082258	AX082258 Sequence
34	336	45.4	777	6	AX536205	AX536205 Sequence
35	336	45.4	777	6	AX536222	AX536222 Sequence
36	336	45.4	803	6	AX118637	AX118637 Sequence
37	333	45.0	701	10	AF310104	AF310104 Mus muscu
38	322	43.5	626	6	AX118621	AX118621 Sequence
39	321	43.4	340	6	AX118622	AX118622 Sequence
40	316.2	42.7	579	6	AR256274	AR256274 Sequence
41	316.2	42.7	579	6	AX082260	AX082260 Sequence
42	316	42.7	582	10	AB053165	AB053165 Rattus no
43	309.8	41.9	321	6	AX408419	AX408419 Sequence
44	293.8	39.7	579	6	AX536207	AX536207 Sequence
45	276.8	37.4	605	6	AX118639	AX118639 Sequence

ALIGNMENTS

RESULT 1
AR256267
LOCUS AR256267 740 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 48 from patent US 6482933.
ACCESSION AR256267
VERSION AR256267.1 GI:27305690
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 740)
AUTHORS Bertin,J.
TITLE Molecules of the card-related protein family and uses thereof
JOURNAL Patent: US 6482933-A 48 19-NOV-2002;
FEATURES Location/Qualifiers


```
QY 721 ATACGAAAGCCAGCTTGAA 740
|||||
Db 721 ATACGAAAGCCAGCTTGAA 740

RESULT 3
AX536208
LOCUS AX536208
DEFINITION Sequence 4 from Patent WO0244354.
ACCESSION AX536208
VERSION AX536208.1 GI:25262599
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE Bertin,J.
AUTHORS Novel molecules of the card-related protein family and uses thereof
TITLE Patent: WO 0244354-A 4 06-JUN-2002;
JOURNAL MILLENIUM PHARMACEUTICALS, INC. (US)
FEATURES
source
location/Qualifiers
1..740
/mol_type="genomic DNA"
/db_xref="taxon:9606"
54..641
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD57246.1"
/db_xref="GI:25262600"
/translation="MGRARDAILDALLENLTABELKKFLKLLSVPLREGYGRIPRGAL
LSMDALDLDKLVSVLYTGAEALTNVLRDMLQEMAGQLQAAATGCGGAPAGIQA
PPQSAKPGHLIDHRAALIAIARTVNEWLLDALYKVLTDDEYQAVRAEPTNPKNMR
KLFSTPAWNWTKDILLQALRESQSYLVEDLERS"
BASE COUNT 146 a 238 c 236 g 120 t
ORIGIN

Query Match 100.0%; Score 740; DB 6; Length 740;
Best Local Similarity 100.0%; Pred. No. 3.1e-107;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGTCCGGCTGACGGGGGTGAGCGGGGATCGCGGGGATCGTGAGGCATGGGCG 60
Db 1 CGCGTCCGGCTGACGGGGGTGAGCGGGGATCGCGGGGATCGTGAGGCATGGGCG 60

QY 61 GCGCGCGGACGCGCATCTCTGGATGCGCTGGAGAACTGACCGCGGAGGAGCTCAAGAGT 120
Db 61 GCGCGCGGACGCGCATCTCTGGATGCGCTGGAGAACTGACCGCGGAGGAGCTCAAGAGT 120

QY 121 TCNAGCTGAAGCTGCTCGGTCCGCTGCGGAGGGCTACGGGCGCATCCCGCGGGCG 180
Db 121 TCNAGCTGAAGCTGCTCGGTCCGCTGCGGAGGGCTACGGGCGCATCCCGCGGGCG 180

QY 181 CGCTGTGTCATGAGCGCTTGGAACCTACCGACAAGCTGGTCAAGCTTCTTACCTGGAGA 240
Db 181 CGCTGTGTCATGAGCGCTTGGAACCTACCGACAAGCTGGTCAAGCTTCTTACCTGGAGA 240

QY 241 CTTACGGCGCGAGCTCACCGCTAACCGTGTGCGGACATGGGCTTCGAGGAGATGGCGG 300
Db 241 CTTACGGCGCGAGCTCACCGCTAACCGTGTGCGGACATGGGCTTCGAGGAGATGGCGG 300

QY 301 GGAGCTGACGGGGCGGACGACACAGGCTCTGAGCGGGCGGCGAGCTGGGATCCAGGCC 360
Db 301 GGAGCTGACGGGGCGGACGACACAGGCTCTGAGCGGGCGGCGAGCTGGGATCCAGGCC 360

QY 361 CTCTCAGTCGGCAGCAAGCGAGGCTGCACCTTTATAGACACAGGCTGGGCTTGA 420
Db 361 CTCTCAGTCGGCAGCAAGCGAGGCTGCACCTTTATAGACACAGGCTGGGCTTGA 420

QY 421 TCCGAGGGGTCAAAACGTTGAGTGGCTGTGGATGCTGTGATCGCGGAAGGTCTTGACGG 480
|||||
```

```
Db 421 TCGCAGGGGTCAAAACGTTGAGTGGCTGTGGATGCTGTGATCGGGAAGGTCTTGACGG 480
QY 481 ATGACGAGTACAGGAGTGGGGCGGAGCCACCAACCAAGCAAGATGCGGAAGCTCT 540
|||||
Db 481 ATGACGAGTACAGGAGTGGGGCGGAGCCACCAACCAAGCAAGATGCGGAAGCTCT 540
QY 541 TCAGTTTACACGAGCCTGGAACTGGACCTGCAAGGACTTGTCTCTCAGGCGCTAAGGG 600
Db 541 TCAGTTTACACGAGCCTGGAACTGGACCTGCAAGGACTTGTCTCTCAGGCGCTAAGGG 600
QY 601 AGTCCAGTCTTACCTGGTGGAGGACCTGGAGCGGAGCTGAGGCTCTTCCAGCAACAC 660
Db 601 AGTCCAGTCTTACCTGGTGGAGGACCTGGAGCGGAGCTGAGGCTCTTCCAGCAACAC 660
QY 661 TCCGCTGACGCCCTGGCAATCCACAAATCATCTGAATCTGATCTTTTATACACAAT 720
Db 661 TCCGCTGACGCCCTGGCAATCCACAAATCATCTGAATCTGATCTTTTATACACAAT 720
QY 721 ATACGAAAGCCAGCTTGAA 740
|||||
Db 721 ATACGAAAGCCAGCTTGAA 740

RESULT 4
AX536223/c
LOCUS AX536223
DEFINITION Sequence 19 from Patent WO0244354.
ACCESSION AX536223
VERSION AX536223.1 GI:25262610
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE Bertin,J.
AUTHORS Novel molecules of the card-related protein family and uses thereof
TITLE Patent: WO 0244354-A 19 06-JUN-2002;
JOURNAL MILLENIUM PHARMACEUTICALS, INC. (US)
FEATURES
source
location/Qualifiers
1..740
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 120 a 236 c 238 g 146 t
ORIGIN

Query Match 100.0%; Score 740; DB 6; Length 740;
Best Local Similarity 100.0%; Pred. No. 3.1e-107;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGTCCGGCTGACGGGGGTGAGCGGGGATCGCGGGGATCGTGAGGCATGGGCG 60
Db 740 CGCGTCCGGCTGACGGGGGTGAGCGGGGATCGCGGGGATCGTGAGGCATGGGCG 681

QY 61 GCGCGCGGACGCGCATCTCTGGATGCGCTGGAGAACTGACCGCGGAGGAGCTCAAGAGT 120
Db 680 GCGCGCGGACGCGCATCTCTGGATGCGCTGGAGAACTGACCGCGGAGGAGCTCAAGAGT 621

QY 121 TCNAGCTGAAGCTGCTCGGTCCGCTGCGGAGGGCTACGGGCGCATCCCGCGGGCG 180
Db 620 TCNAGCTGAAGCTGCTCGGTCCGCTGCGGAGGGCTACGGGCGCATCCCGCGGGCG 561

QY 181 CGCTGTGTCATGAGCGCTTGGAACCTACCGACAAGCTGGTCAAGCTTCTTACCTGGAGA 240
Db 560 CGCTGTGTCATGAGCGCTTGGAACCTACCGACAAGCTGGTCAAGCTTCTTACCTGGAGA 501

QY 241 CTTACGGCGCGAGCTCACCGCTAACCGTGTGCGGACATGGGCTTCGAGGAGATGGCGG 300
Db 500 CTTACGGCGCGAGCTCACCGCTAACCGTGTGCGGACATGGGCTTCGAGGAGATGGCGG 441

QY 301 GGAGCTGACGGGGCGGACGACCAAGGCTCTGGAAGCGCGGAGCTGGGATCCAGGCC 360
|||||
```

Db 440 GGCAGCTGCAGGCGCGCCAGCCAGCGGCTCTGGAGCGCGCAGCTGGGATCCAGGCC 381
QY 361 CTCCTCAGTCGCGAGCCAGCCAGGCTGCACCTTTATAGACCAGCACCGGCTGCGCTTA 420
Db 380 CTCCTCAGTCGCGAGCCAGCCAGGCTGCACCTTTATAGACCAGCACCGGCTGCGCTTA 321
QY 421 TCGCAGGGGTCAAAACGTTGAGTGCTGCTGGATGCTCTGTACGGGAAGTCTCTGACGG 480
Db 320 TCGCAGGGGTCAAAACGTTGAGTGCTGCTGGATGCTCTGTACGGGAAGTCTCTGACGG 261
QY 481 ATGAGCAGTACAGGCGAGTCGCGCGCGAGCCACCAACCAAGCAAGATGCGGAAGCTCT 540
Db 260 ATGAGCAGTACAGGCGAGTCGCGCGCGAGCCACCAACCAAGCAAGATGCGGAAGCTCT 201
QY 541 TCAGTTTACACAGCTGGAACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG 600
Db 200 TCAGTTTACACAGCTGGAACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCT 141
QY 601 AGTCCAGTCTCTACCTGGTGGAGGACCTGGAGCGGAGCTGAGGCTCTCTCCAGCAACAC 660
Db 140 AGTCCAGTCTCTACCTGGTGGAGGACCTGGAGCGGAGCTGAGGCTCTCTCCAGCAACAC 81
QY 661 TCGGTCAGCCCCCTGGCAAATCCCAACCAATCATCTGAAATCTGATCTTTTATACAAAT 720
Db 80 TCGGTCAGCCCCCTGGCAAATCCCAACCAATCATCTGAAATCTGATCTTTTATACAAAT 21
QY 721 ATACGAAAAGCCAGCTTGAA 740
Db 20 ATACGAAAAGCCAGCTTGAA 1

RESULT 5
AP384665 740 bp mRNA linear PRI 19-JUN-2001
LOCUS Homo sapiens caspase recruitment domain protein 5 mRNA, complete cds.
ACCESSION AF384665 GI:14488058
VERSION AF384665.1
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE
AUTHORS Bertin,J.
TITLE CARD5 Protein is a CARD/PYRIN family member that is involved in apoptosis signal transduction
JOURNAL Unpublished
AUTHORS Bertin,J.
TITLE Direct Submission
JOURNAL Submitted (23-MAY-2001) Neurobiology, Millennium Pharmaceuticals Inc., 640 Memorial Drive, Cambridge, MA 02139, USA

FEATURES
Source
1..740
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
54..641
/notes="CARD5; CARD/PYRIN family member; bipartite protein comprised of an N-terminal PYRIN domain and a C-terminal CARD domain; involved in apoptosis signal transduction"
/codon_start=1
/product="caspase recruitment domain protein 5"
/protein_id="AAK63850.1"
/db_xref="GI:14488059"
/translation="MGRARDAITLDALENLTARELKKFKLLSLVPLREGYGRIPRGAL
LSMDALDLTKLVSFYLEYTGAEITANVLRLDMGLQEMAGQLQAATHQSGRAPAGIAQ
PFQSAKPGILGIDHRLAIARVTNWEMLDLALYKGLVLTDEQYQAVRAEPTNPNSMR
KLFSTFPANWTKLLQLALRESQSYLVEDLERS"
146 a 238 c 236 g 120 t

BASE COUNT
ORIGIN

Query Match 100.0%; Score 740; DB 9; Length 740;
Best Local Similarity 100.0%; Pred. No. 3.1e-107;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGCTCCGGCTGCACGCGGGTGAGCGCGGAGCGGCGGAGATCTCTGAGGCATGGGGC 60
Db 1 CCGCTCCGGCTGCACGCGGGTGAGCGCGGAGCGGCGGAGATCTCTGAGGCATGGGGC 60
QY 61 GCGCGGCGAGCGCATCTCTGGATGCGCTGGAGAACTCACCGCCGAGGAGCTCAAGAGT 120
Db 61 GCGCGGCGAGCGCATCTCTGGATGCGCTGGAGAACTCACCGCCGAGGAGCTCAAGAGT 120
QY 121 TCAAGCTGAAGCTGCTGCTCGGTCCGCTGCGCGAGGGCTACGGGCGCATCCCGCGGGCG 180
Db 121 TCAAGCTGAAGCTGCTGCTCGGTCCGCTGCGCGAGGGCTACGGGCGCATCCCGCGGGCG 180
QY 181 CGCTCTGTCCATGACCGCTTTGGACCTCACCGCAAGCTGGTCACTTACCTTGGAGA 240
Db 181 CGCTCTGTCCATGACCGCTTTGGACCTCACCGCAAGCTGGTCACTTACCTTGGAGA 240
QY 241 CCTACGGCGCGAGCTCACCGCTACCGTACGCTGCTGCGCGACATGGGCTCGAGGATGGCG 300
Db 241 CCTACGGCGCGAGCTCACCGCTACCGTACGCTGCTGCGCGACATGGGCTCGAGGATGGCG 300
QY 301 GGCAGCTCGAGCGCGCACGACACAGGCTCTGAGAGCGCGCAGCTGGATCCAGGCC 360
Db 301 GGCAGCTCGAGCGCGCACGACACAGGCTCTGAGAGCGCGCAGCTGGATCCAGGCC 360
QY 361 CTCCTCAGTCGCGAGCCCAAGCCAGGCTGCACTTTATAGACAGCACCGGCTGCGCTTA 420
Db 361 CTCCTCAGTCGCGAGCCCAAGCCAGGCTGCACTTTATAGACAGCACCGGCTGCGCTTA 420
QY 421 TCGCAGGGTCACAAACCTTGAGTGGCTGCTGGATGCTCTGTACGGGAGTCTCTGACGG 480
Db 421 TCGCAGGGTCACAAACCTTGAGTGGCTGCTGGATGCTCTGTACGGGAGTCTCTGACGG 480
QY 481 ATGAGCAGTACAGGCGCGGAGCCAGCCACCAAGCAAGATGCGGAAGCTCT 540
Db 481 ATGAGCAGTACAGGCGCGGAGCCAGCCACCAAGCAAGATGCGGAAGCTCT 540
QY 541 TCAGTTTACACACCGCTGGAACTGGACCTGCAAGGACTTGCTCTCCAGGCCCTTAAGG 600
Db 541 TCAGTTTACACACCGCTGGAACTGGACCTGCAAGGACTTGCTCTCCAGGCCCTTAAGG 600
QY 601 AGTCCAGTCTCTACCTGGTGGAGGACCTGGAGCGGAGCTGAGGCTCTCCAGCAACAC 660
Db 601 AGTCCAGTCTCTACCTGGTGGAGGACCTGGAGCGGAGCTGAGGCTCTCCAGCAACAC 660
QY 661 TCGGTCAGCCCCCTGGCAAATCCCAACCAATCATCTGAAATCTGATCTTTTATACAAAT 720
Db 661 TCGGTCAGCCCCCTGGCAAATCCCAACCAATCATCTGAAATCTGATCTTTTATACAAAT 720
QY 721 ATACGAAAAGCCAGCTTGAA 740
Db 721 ATACGAAAAGCCAGCTTGAA 740

RESULT 6
AX118619 770 bp DNA linear PAT 11-MAY-2001
LOCUS AX118619
DEFINITION Sequence 2 from Patent WO0129235.
ACCESSION AX118619
VERSION AX118619.1 GI:14035570
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Vertino,P.M.
TITLE Tms1 compositions and methods of use
JOURNAL Patent: WO 0129235-A 2 26-APR-2001;
Emory University (US)

RESULT 7	AF184073	770 bp	linear	PRI 02-FBB-2001
LOCUS	AF184073	770 bp	linear	PRI 02-FBB-2001
DEFINITION	Homo sapiens target of methylation-induced silencing 1 (TMS1) mRNA, complete cds.			
ACCESSION	AF184073			
VERSION	AF184073.1	GI:9863863		
KEYWORDS	Homo sapiens (human)			
SOURCE	Homo sapiens			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 770)			
AUTHORS	Conway,K.E., McConnell,B.B., Bowring,C.E., Donald,C.D., Warren,S.T. and Vertino,P.M.			
TITLE	TMS1, a novel proapoptotic caspase recruitment domain protein, is a target of methylation-induced gene silencing in human breast cancers			
JOURNAL	Cancer Res.	60 (22)	6236-6242	(2000)
MEDLINE	20552139			
PUBMED	11103776			
REFERENCE	2 (bases 1 to 770)			
AUTHORS	McConnell,B.B. and Vertino,P.M.			
TITLE	Activation of a caspase-9-mediated apoptotic pathway by subcellular redistribution of the novel caspase recruitment domain protein TMS1			
JOURNAL	Cancer Res.	60 (22)	6243-6247	(2000)
MEDLINE	20552140			
PUBMED	11103777			
REFERENCE	3 (bases 1 to 770)			
AUTHORS	Vertino,P.M.			
TITLE	Direct Submission			
JOURNAL	Submitted (09-SEP-1999) Radiation Oncology, Emory University School of Medicine, 145 Edgewood Avenue, SE, Atlanta, GA 30335, USA			
FEATURES	Location/Qualifiers			
source	1..770			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/clone="IMAGE:159538; IMAGE:1990931"			
	/clone_lib="Soares Breast 3NBHst; NCI_CGAP_Ut3 uterine tumor"			
gene	1..770			
CDS	/gene="TMS1"			
	75..662			
	/gene="TMS1"			
	/function="involved in apoptosis"			
	/note="contains caspase-recruitment domain; CARD protein; target of methylation-mediated gene silencing in Homo sapien tumors"			
	/codon_start=1			
	/product="target of methylation-induced silencing 1"			
	/protein_id="AAG01188.1"			
	/db_xref="GI:9863864"			
	/translation="MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPGAL LMDALDLTKLVSYLETYGAELTANVLRDMGLOEMAGOLQAATHQGSAAAPAGIQ PPSAAKPGLFIDQHEAALIAVRTNVVELLDALYGVLTDEQYQAVRAEPTNPSKWR KLFSTPAWNWTKDLLLQALRESQSVLVEDLERS"			
BASE COUNT	157 a	247 c	241 g	125 t
ORIGIN				
Query Match	99.2%; Score 734; DB 9; Length 770;			
Best Local Similarity	100.0%; Pred. No. 2.8e-106;			
Matches	734;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	7	CGGCTGCACGCGGGTGAGCGCGCAGCGCGCGGGATCTCTGGAGCCATCGGGCGCGCGC	66	
Db	28	CGGCTGCACGCGGGTGAGCGCGCAGCGCGCGGGATCTCTGGAGCCATCGGGCGCGCGC	87	
QY	67	GGCAGCCATCTCGATCGCTGGAGACCTGACCGCGGAGGACTCAAGAAGTTCAAGC	126	
Db	88	GGCAGCCATCTCGATCGCTGGAGACCTGACCGCGGAGGACTCAAGAAGTTCAAGC	147	
QY	127	TGAAGCTGCTGTCGGTGGCGTGCAGGAGGCTACCGGCGCATCCCGGGCGCGCTGC	186	
Db	148	TGAAGCTGCTGTCGGTGGCGTGCAGGAGGCTACCGGCGCATCCCGGGCGCGCTGC	207	
QY	187	TGTCCATGACGCCCTTGACCTCACCGACAAGCTGTGAGTCTTCTACTGAGACCTACG	246	
Db	208	TGTCCATGACGCCCTTGACCTCACCGACAAGCTGTGAGTCTTCTACTGAGACCTACG	267	
QY	247	GGCGGAGCTCACCGCTAACGCTGTCGGCAGCATGGGCTCCAGAGATGCCGGGCGAGC	306	
Db	268	GGCGGAGCTCACCGCTAACGCTGTCGGCAGCATGGGCTCCAGAGATGCCGGGCGAGC	327	
QY	307	TGCAGCGGCCACGACACAGGCTCTGGAGCGCGCCAGCTGGGATCCAGGCCCTCTCTC	366	
Db	328	TGCAGCGGCCACGACACAGGCTCTGGAGCGCGCCAGCTGGGATCCAGGCCCTCTCTC	387	
QY	367	AGTGGCAGCCAAAGCAGGCTTGCACTTTATAGACAGCAGCGGCTGCGCTTATCGGA	426	
Db	388	AGTGGCAGCCAAAGCAGGCTTGCACTTTATAGACAGCAGCGGCTGCGCTTATCGGA	447	
QY	427	GGGTACAAACCTTGAGTGGCTGCTGATGCTCTGACGGAAGTCTCAGCGATCAGC	486	
Db	448	GGGTACAAACCTTGAGTGGCTGCTGATGCTCTGACGGAAGTCTCAGCGATCAGC	507	
QY	487	AGTACAGGACAGTGGCGGCGGAGCCACCAACCCAAAGACAGATCGGAAGCTCTTCAGTT	546	
Db	508	AGTACAGGACAGTGGCGGCGGAGCCACCAACCCAAAGACAGATCGGAAGCTCTTCAGTT	567	
QY	547	TCAACAGCCTGGAACTCGGACCTGCAAGGACTTCTCTCTCAGGCCCTAAGGAGTCCC	606	
Db	568	TCAACAGCCTGGAACTCGGACCTGCAAGGACTTCTCTCTCAGGCCCTAAGGAGTCCC	627	
QY	607	AGTCTTACCTGGTGGAGGACCTGGAGCGGAGCTGAGGCTCTCTCCAGCAACATCCGGT	666	
Db	628	AGTCTTACCTGGTGGAGGACCTGGAGCGGAGCTGAGGCTCTCTCCAGCAACATCCGGT	687	
QY	667	CAGCCCTGGCAATCCCAACAAATCATCTGAAATCTGATCTTTTTTATACAAATATACGA	726	
Db	688	CAGCCCTGGCAATCCCAACAAATCATCTGAAATCTGATCTTTTTTATACAAATATACGA	747	
QY	727	AAAGCCAGCTTGA	740	
Db	748	AAAGCCAGCTTGA	761	

Qy	127	TGAAGCTGCTCTCGGTCCGCTGCGGAGAGGCTACGGGGCGATCCGGGGGGCGCGCTGC	188		
Db	148	TGAAGCTGCTCTCGGTGCGCTGCGGAGAGGCTACGGGGCGATCCGGGGGGCGCGCTGC	207		
Qy	187	TGTCATGGAGCGCTTGGACCTCACCGACAAGCTGCTCAGCTTCTACCTTGGAGACCTACG	246		
Db	208	TGTCCATGGAGCGCTTGGACCTCACCGACAAGCTGCTCAGCTTCTACCTTGGAGACCTACG	267		
Qy	247	GCGCCGAGCTCACCGCTAAACGTGCTCGCGCATGGGCTGACGAGATGGCCGGGCGAGC	306		
Db	268	GCGCCGAGCTCACCGCTAAACGTGCTCGCGCATGGGCTGACGAGATGGCCGGGCGAGC	327		
Qy	307	TGCAGGCGGCCACGACACGAGGCTCTGAGCGCGCCAGCTGGGATCCAGGCCCCCTCTC	366		
Db	328	TGCAGGCGGCCACGACACGAGGCTCTGAGCGCGCCAGCTGGGATCCAGGCCCCCTCTC	387		
Qy	367	AGTCGGGAGCAAGCCAGGCTGCACTTTATAGACCAGCACGGGCTGGCTTATCGCGA	426		
Db	388	AGTCGGGAGCAAGCCAGGCTGCACTTTATAGACCAGCACGGGCTGGCTTATCGCGA	447		
Qy	427	GGGTCAACAAAGCTTGAGTGGCTGCTGACGGGAAAGTCTCTGACGGATGAGC	486		
Db	448	GGGTCAACAAAGCTTGAGTGGCTGCTGAGATGCTGTACGGGAAAGTCTCTGACGGATGAGC	507		
Qy	487	AGTACCAGGCAAGTCGGGCGCCAGGCCACCAACCCAAAGCAAGATGCGGAAGCTCTTCAGTT	546		
Db	508	AGTACCAGGCAAGTCGGGCGCCAGGCCACCAACCCAAAGCAAGATGCGGAAGCTCTTCAGTT	567		
Qy	547	TCACACGAGCTGNACTGGACCTGCAGGACTTGCTCTCCAGGCGCTTAAAGGAGTCCC	606		
Db	568	TCACACGAGCTGNACTGGACCTGCAGGACTTGCTCTCCAGGCGCTTAAAGGAGTCCC	627		
Qy	607	AGTCCTACTCGTGGAGGACCTGGAGCGGAGCTGAGGCTCTCTCCAGCAACACTCCCGT	666		
Db	628	AGTCCTACTCGTGGAGGACCTGGAGCGGAGCTGAGGCTCTCTCCAGCAACACTCCCGT	687		
Qy	667	CAGCCCTGGCAATCCCAACCAATCATCTGAATCTGATCTTTTTATACAAATATACGA	726		
Db	688	CAGCCCTGGCAATCCCAACCAATCATCTGAATCTGATCTTTTTATACAAATATACGA	747		
Qy	727	AAAGCCAGCTTGAA	740		
Db	748	AAAGCCAGCTTGAA	761		
RESULT 8	AX459863	782 bp	DNA	linear	PAT 08-JUL-2000
LOCUS	Sequence 6 from Patent WO240668.				
DEFINITION	AX459863				
ACCESSION	AX459863.1	GI:21725640			
VERSION					
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1				
AUTHORS	Tschopp, J. and Martinon, F.				
TITLE	Proteins and dna sequences underlying these proteins used for treating inflammations				
JOURNAL	Patent: WO 0240668-A 6 23-MAY-2002;				
FEATURES	Apotech Research and Development Ltd. (CH)				
source	Location/Qualifiers				
	1..782				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
	/note="Pycard.cdna"				
BASE COUNT	158 a 251 C 246 G				127 t
ORIGIN					

Query Match

Best Local Similarity 99.2%; Score 734; DB 6; Length 782;

100.0%; Pred. No. 2.8e-106;

MEDLINE 20036508
PUBMED 10567338
REFERENCE 2 (bases 1 to 782)
AUTHORS Masumoto,J., Sagara,J. and Taniguchi,S.
TITLE Direct Submission
JOURNAL Submitted (04-FEB-1999) Junya Masumoto, Shinsu University School
of Medicine, Research Center on Aging and Adaptation; Asahi 3-1-1,
Matsumoto, Nagano 390-8621, Japan
(E-mail:masumoto@sch.md.shinsu-u.ac.jp, Tel:81-263-37-2723,
Fax:81-263-37-2724)
COMMENT On Oct 14, 2000 this sequence version replaced gi:6482371.
SEQUENCE Sequence updated (29-Nov-1999).
FEATURES Location/Qualifiers
 1..782
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="16"
 /map="16p11.2-12"
 1..782
 /genes="ASC"
 87..674
 /genes="ASC"
 /codon_start=1
 /product="apoptosis-associated speck-like protein
 containing a CARD"
 /protein_id="BAA87339.2"
 /db_xref="GI:10801602"
 /translation="MGRARDAILDALENITABELKKFKLLSVPLRBYGRIPRGAL
 LSMDALDTDLKLVSYLETYTGAEITANVLRLDGLQEMAGQQAATHQSGGAAPAGIQ
 PPQSAKPGPLHIDQHRALAIARVNTNWLWLLDALYKVLTDQYQAVRAEPTNPMSKMR
 KLFTPTPAWNTCKOLLQALRESQSYLVEDLERS"
BASE COUNT 158 a 251 c 246 g 127 t
ORIGIN
Query Match 99.2%; Score 734; DB 9; Length 782;
Best Local Similarity 100.0%; Pred. No. 2.8e-106;
Matches 734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 CGGCTGCAGCGGGTGAGCGCGGCGGCGGCGGATCCTGGAGCATGGGGCGGCGC 66
DB 40 CGGCTGCAGCGGGTGAGCGCGGCGGCGGCGGATCCTGGAGCATGGGGCGGCGC 99
QY 67 GCGACGCCATCTGGATGCGTGGAGAACTGTACCGCGGAGAGCTCAAGAAGTTCAAGC 126
DB 100 GCGACGCCATCTGGATGCGTGGAGAACTGTACCGCGGAGAGCTCAAGAAGTTCAAGC 159
QY 127 TGAAGCTGCTGTGCGTCCGCTGCGCGAGGGCTACGGGCGCATCCCGCGGGCGGCGTCC 186
DB 160 TGAAGCTGCTGTGCGTCCGCTGCGCGAGGGCTACGGGCGCATCCCGCGGGCGGCTGC 219
QY 187 TGTTCATGAGCGCTTGGACCTCACCAGAACAGTGTGTACGTTTACCTGGAGACCTACG 246
DB 220 TGTTCATGAGCGCTTGGACCTCACCAGAACAGTGTGTACGTTTACCTGGAGACCTACG 279
QY 247 GCGCCGAGCTCACCGCTACGCTGCGCGACATGGGCTGCGAGGAGATGGCGGGCAGC 306
DB 280 GCGCCGAGCTCACCGCTACGCTGCGCGACATGGGCTGCGAGGAGATGGCGGGCAGC 339
QY 307 TGCAGCGGCGCACCGACCGGGCTCTGGAGCGCGCGAGCTGGGATCCAGGCCCTTCCTC 366
DB 340 TGCAGCGGCGCACCGACCGGGCTCTGGAGCGCGCGAGCTGGGATCCAGGCCCTTCCTC 399
QY 367 AGTTCGAGCGCAAGCGAGCGCTGCACTTTATAGACAGAGCCGGGTGCGCTTATCGCGA 426
DB 400 AGTTCGAGCGCAAGCGAGCGCTGCACTTTATAGACAGAGCCGGGTGCGCTTATCGCGA 459
QY 427 GGTTCACAACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 486
DB 460 GGTTCACAACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 519
QY 487 AGTACAGGCGAGTGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 546

DB 520 AGTACCAGGCGAGTGGGGCGGAGCCCAACCAAGCAAGATGCGGAAGCTCTTCAGTT 579
QY 547 TCACACCAAGCTGGAACCTGGACCTGCAAGGACTTGTCTCTCCAGGCCCTTAAGGGAGTCCC 606
DB 580 TCACACCAAGCTGGAACCTGGACCTGCAAGGACTTGTCTCTCCAGGCCCTTAAGGGAGTCCC 639
QY 607 AGTCTTACTGCTGGTAGGAGCACTGAGCGGAGCTGAGGCTCTCTCCAGCAACACTCCGGT 666
DB 640 AGTCTTACTGCTGGTAGGAGCACTGAGCGGAGCTGAGGCTCTCTCCAGCAACACTCCGGT 699
QY 667 CAGCCCTTGGCAATCCCAACCAATCATCTGAATCTTTTATACAAATATACGA 726
DB 700 CAGCCCTTGGCAATCCCAACCAATCATCTGAATCTTTTATACAAATATACGA 759
QY 727 AAAGCCAGCTTGAA 740
DB 760 AAAGCCAGCTTGAA 773
RESULT 10
AX017270
LOCUS AX017270 779 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 21 from Patent WO9947669.
ACCESSION AX017270
VERSION AX017270.1 GI:10042188
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and
Pilarczyk,C.
TITLE Human nucleic acid sequences from tissue of breast tumors
JOURNAL Patent: WO 9947669-A 21 23-SEP-1999;
SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN
(DE); PILARSKY CHRISTIAN (DE)
FEATURES Location/Qualifiers
 1..779
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
BASE COUNT 149 a 252 c 251 g 127 t
ORIGIN
Query Match 99.1%; Score 733; DB 6; Length 779;
Best Local Similarity 100.0%; Pred. No. 4e-106;
Matches 733; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 CGGCTGCAGCGGGTGAGCGCGGCGGCGGATCCTGGAGCATGGGGCGGCGC 66
DB 47 CGGCTGCAGCGGGTGAGCGCGGCGGCGGATCCTGGAGCATGGGGCGGCGC 106
QY 67 GCGACGCCATCTGGATGCGTGGAGAACTGTACCGCGGAGAGCTCAAGAAGTTCAAGC 126
DB 107 GCGACGCCATCTGGATGCGTGGAGAACTGTACCGCGGAGAGCTCAAGAAGTTCAAGC 166
QY 127 TGAAGCTGCTGTGCGTCCGCTGCGCGAGGGCTACGGGCGCATCCCGCGGGCGGCTGC 186
DB 167 TGAAGCTGCTGTGCGTCCGCTGCGCGAGGGCTACGGGCGCATCCCGCGGGCGGCTGC 226
QY 187 TGTTCATGAGCGCTTGGACCTCACCAGAACAGTGTGTACGTTTACCTGGAGACCTACG 246
DB 227 TGTTCATGAGCGCTTGGACCTCACCAGAACAGTGTGTACGTTTACCTGGAGACCTACG 286
QY 247 GCGCCGAGCTCACCGCTAACGCTGCGCGACATGGGCTGCGAGGAGATGGCGGGCAGC 306
DB 287 GCGCCGAGCTCACCGCTAACGCTGCGCGACATGGGCTGCGAGGAGATGGCGGGCAGC 346
QY 307 TGCAGCGGCGCACCGACCGAGGCTCTGGAGCGCGCGAGCTGGGATCCAGGCCCTTCCTC 366
DB 347 TGCAGCGGCGCACCGACCGAGGCTCTGGAGCGCGCGAGCTGGGATCCAGGCCCTTCCTC 406


```
source 1..779
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606" 127 t
BASE COUNT 149 a 252 c 251 g 127 t
ORIGIN
Query Match 99.1%; Score 733; DB 6; Length 779;
Best Local Similarity 100.0%; Pred. No. 4e-106;
Matches 733; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 CGGCTGCAGCGGGTGCAGCGCGGGGAGCTCTCGAGCCATGGGGCGGCGC 66
DB 47 CGGCTGCAGCGGGTGCAGCGCGGGGAGCTCTCGAGCCATGGGGCGGCGC 106
QY 67 GCGAGCGCATCTGATGCTGAGAGACCTGACCGCGGAGGAGCTCAAGAGTTCAAGC 126
DB 107 GCGAGCGCATCTGATGCTGAGAGACCTGACCGCGGAGGAGCTCAAGAGTTCAAGC 166
QY 127 TGAAGCTGCTGTCGGTGCCTGCGAGAGGCTACGGGCGCATCCGCGGGGCGGCTGC 186
DB 167 TGAAGCTGCTGTCGGTGCCTGCGAGAGGCTACGGGCGCATCCGCGGGGCGGCTGC 226
QY 187 TGTCCATGACCGCTTGGACCTTACCGACAAGCTGTCAGCTTCTACCTGGAGACTAGC 246
DB 227 TGTCCATGACCGCTTGGACCTTACCGACAAGCTGTCAGCTTCTACCTGGAGACTAGC 286
QY 247 GCGCGAGCTCACCGCTTAACGCTGTCGCGGACATGSGGCTGCGAGAGATGGCGGGCAGC 306
DB 287 GCGCGAGCTCACCGCTTAACGCTGTCGCGGACATGSGGCTGCGAGAGATGGCGGGCAGC 346
QY 307 TGCAGCGCGCCACGACCAAGGCTCTGGAGCGCGCTGAGCTGGATGCCAGGCCCTCTCTC 366
DB 347 TGCAGCGCGCCACGACCAAGGCTCTGGAGCGCGCTGAGCTGGATGCCAGGCCCTCTCTC 406
QY 367 AGTTCGAGCGCAAGCCAGCGCTGCACTTTATAGACAGCAACCGGGTGGCTTTATCGCGA 426
DB 407 AGTTCGAGCGCAAGCCAGCGCTGCACTTTATAGACAGCAACCGGGTGGCTTTATCGCGA 466
QY 427 GGGTCAACAACTGAGTGGCTGCTGGATGCTCTGTAACGAGGCTCTTACCGAGATGAGC 486
DB 467 GGGTCAACAACTGAGTGGCTGCTGGATGCTCTGTAACGAGGCTCTTACCGAGATGAGC 526
QY 487 AGTACAGGAGCTGCGGGCGAGCCACCAACCCAGCAAGATGCGGAAGCTCTTCAAGTT 546
DB 527 AGTACAGGAGCTGCGGGCGAGCCACCAACCCAGCAAGATGCGGAAGCTCTTCAAGTT 586
QY 547 TCACACGAGCTGGAACTGGAACCTGCAAGGACTTGTCTCTCCAGGCGCTTAAGGAGTCCC 606
DB 587 TCACACGAGCTGGAACTGGAACCTGCAAGGACTTGTCTCTCCAGGCGCTTAAGGAGTCCC 646
QY 607 AGTTCCTACCTGGTGGAGGAGCTGGAGCGAGCTGAGGCTCTTCCAGCAACACTCCGCT 666
DB 647 AGTTCCTACCTGGTGGAGGAGCTGGAGCGAGCTGAGGCTCTTCCAGCAACACTCCGCT 706
QY 667 CAGCCCTGGCAATCCCAACAAATCACTCTGAATCTGATCTTTTATACAAATATACGA 726
DB 707 CAGCCCTGGCAATCCCAACAAATCACTCTGAATCTGATCTTTTATACAAATATACGA 766
QY 727 AAAGCCAGCTTGA 739
DB 767 AAAGCCAGCTTGA 779
```

```
RESULT 13
BD057255
LOCUS BD057255 785 bp DNA linear PAT 27-AUG-2002
DEFINITION Apoptosis-related protein, its antibody and its DNA.
ACCESSION BD057255
VERSION BD057255.1 GI:22602861
KEYWORDS JP 2001275681-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
```

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 785)
AUTHORS Masumoto,J., Sagara,J. and Taniguchi,S.
TITLE Apoptosis-related protein, its antibody and its DNA
JOURNAL Patent: JP 2001275681-A 1 09-OCT-2001;
JUNYA MASUMOTO, JUNJI ARA, SHUNICHIRO TANIGUCHI, MEDICAL &
BIOLOGICAL LABORATORIES CO LTD
COMMENT OS Homo sapiens (human)
PN JP 2001275681-A/1
PD 09-OCT-2001
PF 31-MAR-2000 JP 2000098204
PI JUNYA MASUMOTO, JUNJI SAGARA, SHUNICHIRO TANIGUCHI PC
C12N15/09, C07K14/82, C07K16/32//C12P21/02, C12P21/08, C12N15/00 CC
FH Key Location/Qualifiers.
FEATURES
source 1..785
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 160 a 252 c 247 g 126 t
ORIGIN
Query Match 98.8%; Score 730.8; DB 6; Length 785;
Best Local Similarity 99.7%; Pred. No. 8.8e-106;
Matches 732; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 7 CGGCTGCAGCGGGTGCAGCGCGGGATCTCTGAGCCATGGGGCGGCGC 66
DB 40 CGGCTGCAGCGGGTGCAGCGCGGGATCTCTGAGCCATGGGGCGGCGC 99
QY 67 GCGAGCGCATCTGATGCTGAGAGAACCTGACCGCGGAGGAGCTCAAGAGTTCAAGC 126
DB 100 GCGAGCGCATCTGATGCTGAGAGAACCTGACCGCGGAGGAGCTCAAGAGTTCAAGC 159
QY 127 TGAAGCTGCTGTCGGTGCCTGCGAGAGGCTACGGGCGCATCCGCGGGGCGGCTGC 186
DB 160 TGAAGCTGCTGTCGGTGCCTGCGAGAGGCTACGGGCGCATCCGCGGGGCGGCTGC 219
QY 187 TGTCCATGAGCGCTTGGACCTCACCGACAAGCTGCTGAGCTTCTACCTGGAGACTAGC 246
DB 220 TGTCCATGAGCGCTTGGACCTCACCGACAAGCTGCTGAGCTTCTACCTGGAGACTAGC 279
QY 247 GCGCGAGCTCACCGCTTAACGCTGTCGCGACATGGGCTTCAGGAGATGGCGGGCAGC 306
DB 280 GCGCGAGCTCACCGCTTAACGCTGTCGCGACATGGGCTTCAGGAGATGGCGGGCAGC 339
QY 307 TGCAGCGGCGCAGCAACAGGGCTCTGGAGCGCGGCGAGCTGGGATCCAGGCGCTCTC 366
DB 340 TGCAGCGGCGCAGCAACAGGGCTCTGGAGCGCGGCGAGCTGGGATCCGAGGCTCTC 399
QY 367 AGTTCGAGCGCAAGCCAGCGCTGCACTTTATAGACAGCAACCGGGCTGCTTATCGCGA 426
DB 400 AGTTCGAGCGCAAGCCAGCGCTGCACTTTATAGACAGCAACCGGGCTGCTTATCGCGA 459
QY 427 GGGTCAACAACTGAGTGGCTGCTGGATGCTCTGTAACGGAAGTCTCTGACGAGTACG 486
DB 460 GGGTCAACAACTGAGTGGCTGCTGGATGCTCTGTAACGGAAGTCTCTGACGAGTACG 519
QY 487 AGTACCGAGCTGCGGGCGAGCCCAACCCAGCAAGATGCGGAAGCTCTTCAAGTT 546
DB 520 AGTACCGAGCTGCGGGCGAGCCCAACCCAGCAAGATGCGGAAGCTCTTCAAGTT 579
QY 547 TCACACGAGCTGGAACTGGACCTGCAAGGACTTGTCTCTCCAGGCGCTTAAGGAGTCCC 606
DB 580 TCACACGAGCTGGAACTGGACCTGCAAGGACTTGTCTCTCCAGGCGCTTAAGGAGTCCC 639
QY 607 AGTTCCTACCTGGTGGAGGAGCTGGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGCT 666
DB 640 AGTTCCTACCTGGTGGAGGAGCTGGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGCT 699
QY 667 CAGCCCTGGCAATCCCAACAAATCATCTGAACTCTGATCTTTTATACAAATATACGA 726
```

```

Db      700 CAGCCCTGGCAATCCACCACCAATCATCTGAATCTGATCTTTTATACAAATATACGA 759
QY      727 AAAGCCAGCTTGAA 740
Db      760 AAAGCCAGCTTGAA 773

RESULT 14
AK000211
LOCUS   Homo sapiens cDNA FLJ20204 fis, clone COLF1505.          linear      PRI 22-FEB-2000
DEFINITION
ACCESSION AK000211
VERSION   AK000211.1 GI:7020146
KEYWORDS  oligo capping; fis (full insert sequence).
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE
AUTHORS   Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
            Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
            Nakamura,Y., Isogai,T. and Sugano,S.
            NEDO human cDNA sequencing project
            Unpublished
            2 (bases 1 to 740)
AUTHORS   Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
            Shibahara,T., Tanaka,T. and Nakamura,Y.
            Direct Submission
            Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,
            University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1,
            Minato-ku, Tokyo 108-8639, Japan (E-mail:cdnal@ims.u-tokyo.ac.jp,
            Tel:81-3-5449-5286, Fax:81-3-5449-5416)
COMMENT   NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan; cDNA full insert
            sequencing: Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing: Department of
            Virology and Human Genome Center, Institute of Medical Science,
            University of Tokyo (partly supported by Science and Technology
            Agency).
FEATURES             Location/Qualifiers
     source           1..740
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="COLF1505"
                     /tissue_type="colon mucosa"
                     /clone_lib="ColF"
                     /note="cloning vector pME18SFL3"
     CDS               177..626
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAA91012.1"
                     /db_xref="GI:7020147"
                     /translation="MDALDLDKLVSFYLETYCAELTNVLRDMLQEMAGLOLAATH
                     QSGAAGPAGIOAPPOSAKPLHFIHQRAALIARVNVWLLDLYLKGVLTDQYQA
                     VRAEPTNPCKMRKLFSETPAWNWTCKDLLQLALRESQSVLYVEDLRS"
BASE COUNT          161 a 231 c 230 g 118 t
ORIGIN
Query Match          96.4%; Score 713; DB 9; Length 740;
Best Local Similarity 99.7%; Pred. No. 5.7e-103;
Matches 725; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY      14  ACGGGGTGAGCGCGGACGCGCGGGGATCTCTGGAGCCATGGCGCGCGGACGC 73
Db      1  ACGGGGGTGAAGCGCGGACGCGCGGGGATCTCTGGAGCCATGGG--GCGCGGCGACGC 58
QY      74  CATCTCTGATCGCTGGAGAACTGACCGCCGAGGAGCTCAAGAAGTTCAAGCTGAAGCT 133
Db      59  CATCTCTGATCGCTGGAGAACTGACCGCCGAGGAGCTCAAGAAGTTCAAGCTGAAGCT 118
QY      134 GCTGTGCGTCCGCTGCGGAGGGCTACGGGCGCATCCCGCGGGCGCGCTGCTGTCCAT 193

```

```

Db      119 GCTGTGCTGCGCTGCGCGAGGGGTACGGGGGCATCCCGGGGGCGCTGCTGTCCAT 178
QY      194 GGACGCTTTGACCTTACCCGACCAAGCTGCTCAGCTTTCTACCTTGAGAGACTACGGGCGCGGA 253
Db      179 GGACGCTTTGACCTTACCCGACCAAGCTGCTCAGCTTTCTACCTTGAGAGACTACGGGCGCGGA 238
QY      254 GCTACCCCTAACGCTGCTGCGCGACATGGGCTCTGAGAGATGGCCGGGAGCTGCAGGC 313
Db      239 GCTACCCCTAACGCTGCTGCGCGACATGGGCTCTGAGAGATGGCCGGGAGCTGCAGGC 298
QY      314 GGCAACGACCAAGGGCTCTGAGCGCGCCAGCTGGGATCCAGGCCCTCTCTCAGTCCGC 373
Db      299 GGCAACGACCAAGGGCTCTGAGCGCGCCAGCTGGGATCCAGGCCCTCTCTCAGTCCGC 358
QY      374 AGCCAAGCCAGGCTGTCACCTTTATAGACAGCAGCCGGCTGCGCTTATCGGAGGCTCAC 433
Db      359 AGCCAAGCCAGGCTGTCACCTTTATAGACAGCAGCCGGCTGCGCTTATCGGAGGCTCAC 418
QY      434 AAACGTTGAGTGGCTGCTGGATGCTCTGTACGGGAAGTCTCTGACGGATGAGCAGTACCA 493
Db      419 AAACGTTGAGTGGCTGCTGGATGCTCTGTACGGGAAGTCTCTGACGGATGAGCAGTACCA 478
QY      494 GCGAGTGGCGGCGAGCCCAACCAAGCAAGATGCGGAAGCTCTTCAGTTTCACACC 553
Db      479 GCGAGTGGCGGCGAGCCCAACCAAGCAAGATGCGGAAGCTCTTCAGTTTCACACC 538
QY      554 AGCCTGGAACCTGGACCTGCAAGGACTTCTCTCCAGGCCCTAAGGGAGTCCCACTCTA 613
Db      539 AGCCTGGAACCTGGACCTGCAAGGACTTCTCTCCAGGCCCTAAGGGAGTCCCACTCTA 598
QY      614 CTTGTGGAGGACCTTGGAGCGGAGCTGAGGCTCTTCCAGCAACACCTCCGCTCAGCCCC 673
Db      599 CTTGTGGAGGACCTTGGAGCGGAGCTGAGGCTCTTCCAGCAACACCTCCGCTCAGCCCC 658
QY      674 TGGCAATCCCAACCAATCATCTGATCTGATCTTTTATACAAATATACAAAGCCA 733
Db      659 TGGCAATCCCAACCAATCATCTGATCTGATCTTTTATACAAATATACAAAGCCA 718
QY      734 GCTTGAA 740
Db      719 GCTTGAA 725

RESULT 15
AF310103
LOCUS   Homo sapiens PYCARD mRNA, complete cds.          linear      PRI 04-MAY-2001
DEFINITION
ACCESSION AF310103
VERSION   AF310103.1 GI:11096298
KEYWORDS
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 712)
AUTHORS   Martinon,F., Hofmann,K. and Tschopp,J.
            The pyrin domain: a possible member of the death domain-fold family
            implicated in apoptosis and inflammation
            Curr. Biol. 11 (4), R118-R120 (2001)
JOURNAL   MEDLINE
PUBMED    11250163
REFERENCE 2 (bases 1 to 712)
AUTHORS   Martinon,F., Hofmann,K. and Tschopp,J.
            Direct Submission
            Submitted (28-SEP-2000) Institute of Biochemistry, University of
            Lausanne, Ch des Boveresses 155, Epalinges 1066, Switzerland
JOURNAL
FEATURES             Location/Qualifiers
     source           1..712
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
     CDS               1..588
                     /note="contains PYD and CARD domains"

```

```
/codon_start=1
/product="PYCARD"
/protein_id="AAG30286.1"
/db_xref="GI:11096299"
/translation="MGRARDAILDALENLTAEELKFKLKLLSVPLREGYGRIRGAL
LSMDALDITDKLVSYLETYNELTANVLRDMGLQEMAGQLOATHQSGAAPAGIQ
PPQSAKPGDLHFIDQRRALFARVNVNVEMLDLYKVLTDQYQVRAEPTNPKMR
KLFSPTAWNTCKDLQLQALRESQSYLVIEDLRS"
BASE COUNT      166 a      221 c      210 g      115 t
ORIGIN

Query Match      92.8%; Score 687; DB 9; Length 712;
Best Local Similarity 100.0%; Pred. No. 7.2e-99;
Matches 687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 ATGGGGCGCGCGCGAGCCCATCTGGATGCGCTGGAGAACCTACCGCGGAGGCTC 113
Db 1 ATGGGGCGCGCGCGAGCCCATCTGGATGCGCTGGAGAACCTACCGCGGAGGCTC 60
QY 114 AAGAAATTCAAGCTGAAGTGTGTGCTGCGCTGCGGAGGGCTACGGGCGCATCCCG 173
Db 61 AAGAAATTCAAGCTGAAGTGTGTGCTGCGCTGCGGAGGGCTACGGGCGCATCCCG 120
QY 174 CGGGCGCGCTCTGTTCATGAGCGCTTGGACCTCACCGACAAGTGTGAGTTCTTAC 233
Db 121 CGGGCGCGCTCTGTTCATGAGCGCTTGGACCTCACCGACAAGTGTGAGTTCTTAC 180
QY 234 CTGGAGACCTACGGCGCGAGCTCACCGCTACCTGCTGCGGAGGCTACGGGCGCATCCCG 293
Db 181 CTGGAGACCTACGGCGCGAGCTCACCGCTACCTGCTGCGGAGGCTACGGGCGCATCCCG 240
QY 294 ATGGCGGGGAGCTGCAGCGCGCCACCGACACAGGGCTCTGGAGCGCGCGAGCTGGGATC 353
Db 241 ATGGCGGGGAGCTGCAGCGCGCCACCGACACAGGGCTCTGGAGCGCGCGAGCTGGGATC 300
QY 354 CAGGCGCCCTCTCAGTTCGCGAGCCAAAGCCAGCGCTGCACTTTATAGACAGCAACCGGGCT 413
Db 301 CAGGCGCCCTCTCAGTTCGCGAGCCAAAGCCAGCGCTGCACTTTATAGACAGCAACCGGGCT 360
QY 414 GCGCTTATCGCGAGGGTCAAAACGTTGAGTGGCTGCTGATGCTGTACGGGAGGTC 473
Db 361 GCGCTTATCGCGAGGGTCAAAACGTTGAGTGGCTGCTGATGCTGTACGGGAGGTC 420
QY 474 CTGACGGATGAGCAGTACAGGCGAGTGCAGGGCGAGCCCAACCAAGCAAGATGCGG 533
Db 421 CTGACGGATGAGCAGTACAGGCGAGTGCAGGGCGAGCCCAACCAAGCAAGATGCGG 480
QY 534 AAGCTCTTCAGTTTCAACACGCTGGAACTGGAACCTGAAGACTTGTCTCTCAGGGCC 593
Db 481 AAGCTCTTCAGTTTCAACACGCTGGAACTGGAACCTGAAGACTTGTCTCTCAGGGCC 540
QY 594 CTAAGGGAGTCCAGTCTCTACTGTGGAGGACCTGGAGGGAGCTGAGGCTCTTCCCA 653
Db 541 CTAAGGGAGTCCAGTCTCTACTGTGGAGGACCTGGAGGGAGCTGAGGCTCTTCCCA 600
QY 654 GCAACACTCCGGTACGCGCTGGCAATCCACCAATCATCTGATCTGATCTTTTAT 713
Db 601 GCAACACTCCGGTACGCGCTGGCAATCCACCAATCATCTGATCTGATCTTTTAT 660
QY 714 ACACAATATACGAAAGCCAGCTTGAA 740
Db 661 ACACAATATACGAAAGCCAGCTTGAA 687
```


	1	740	100.0	740	22	AAF30007	Human CARD-5 cDNA.
C	2	740	100.0	740	24	ABK87966	Human caspase recr
	3	740	100.0	740	24	ABK87967	Human caspase recr
	4	740	100.0	740	25	ABSS6032	CDNA encoding huma
	5	734	99.2	770	22	AA003890	Human target of me
	6	734	99.2	782	24	AAL47126	Pyrim domain conta
	7	733	99.1	745	22	AAS59817	Human novel cytoke
	8	733	99.1	779	20	AAZ33631	Human breast tumou
	9	733	99.1	779	20	AAZ33631	Human breast tumou

XX (MILL-) MILLENNIUM PHARM INC.
 PA Bertin J;
 XX
 XX
 XX
 DR WPI; 2001-061973/07.
 DR P-PSDB; AAB20085.
 XX
 XX Isolated intracellular proteins predicted to be involved in regulating
 PT caspase activation are used for diagnosis and treatment of e.g. cancer,
 PT viral infections, autoimmune diseases, neurological diseases and
 PT haematological disorders -
 XX
 PS Claim 1(a); Fig 21; 208pp; English.
 XX
 CC The present sequence is that of cDNA encoding human caspase
 CC recruitment domain 5 (CARD-5, see AAB20085). The cDNA was isolated
 CC from a testis cDNA library using murine CARD-L. Plasmid EphC5
 CC containing CARD-5 cDNA is deposited as ATCC PTA-213. CARD-5
 CC is an intracellular protein predicted to be involved in regulating
 CC caspase activation. It is useful as a modulating agent in
 CC regulating cellular processes include cell growth and cell death.
 CC Methods of diagnosing and treating patients suffering from a
 CC disorder associated with an abnormal level or rate of apoptotic
 CC cell death, abnormal activity of the Fas/APO-1 receptor complex,
 CC abnormal activity of the tumour necrosis factor receptor complex
 CC or abnormal activity of a caspase involve administering a compound
 CC that modulates the expression or activity of CARD-3, CARD-4, CARD-5
 CC or CARD-6 e.g. a small molecule, antisense nucleic acid, ribozyme
 CC or polypeptide. Such disorders include cancer, viral infection,
 CC autoimmune disorders, neurological diseases, haematological
 CC disorders, inflammatory disorders and immune disorders. CARD
 CC nucleic acids can be used to express CARD proteins in a host cell
 CC e.g. for gene therapy applications, to detect a genetic lesion and
 CC to modulate CARD activity.
 XX
 SQ Sequence 740 BP; 146 A; 238 C; 236 G; 120 T; 0 other;
 Query Match 100.0%; Score 740; DB 22; Length 740;
 Best Local Similarity 100.0%; Pred. No. 1.7e-147;
 Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGCGTCCGCTGACGCGGGTGTAGCGCGGCGAGCGCGGGATCTCGAGCCATGGGCG 60
 DB 1 CGCGTCCGCTGACGCGGGTGTAGCGCGGCGAGCGCGGGATCTCGAGCCATGGGCG 60
 QY 61 GCGCGCGCGACGCCATCTCTGGATGGCTGAGAACTGACCGCGAGGAGCTCAAGAACT 120
 DB 61 GCGCGCGCGACGCCATCTCTGGATGGCTGAGAACTGACCGCGAGGAGCTCAAGAACT 120
 QY 121 TCAAGCTGAAGCTGTCTGCTGCGTCCGCTGCGGAGGGCTACCGGCGCATCCCGCGGGCG 180
 DB 121 TCAAGCTGAAGCTGTCTGCTGCGTCCGCTGCGGAGGGCTACCGGCGCATCCCGCGGGCG 180
 QY 181 CGCTGCTGCTATGAGCGCTTGGACCTCACCAGCAAGCTGTCTACCTTCTACCTGGAGA 240
 DB 181 CGCTGCTGCTATGAGCGCTTGGACCTCACCAGCAAGCTGTCTACCTTCTACCTGGAGA 240
 QY 241 CCTACGGCGCGAGCTCACCGCTAACTGCTGCGGACATGGGCGCTGACAGGAGATGGCGG 300
 DB 241 CCTACGGCGCGAGCTCACCGCTAACTGCTGCGGACATGGGCGCTGACAGGAGATGGCGG 300
 QY 301 GGCAGCTGAGGGGCGACGACCGAGGCTCTGGAGCGCGCGAGTGGGATCCAGGGCC 360
 DB 301 GGCAGCTGAGGGGCGACGACCGAGGCTCTGGAGCGCGCGAGTGGGATCCAGGGCC 360
 QY 361 CTCCTCAGTCGCGCAGCAAGCCAGGCTGCACCTTTATAGACAGCAGCGGCTGGCTTA 420
 DB 361 CTCCTCAGTCGCGCAGCAAGCCAGGCTGCACCTTTATAGACAGCAGCGGCTGGCTTA 420
 QY 421 TCGCGAGGGTCAAAACGTTGAGTGGCTGCTGGATGCTCTGTACGGGAGGCTCTGACGG 480
 DB 421 TCGCGAGGGTCAAAACGTTGAGTGGCTGCTGGATGCTCTGTACGGGAGGCTCTGACGG 480

QY 481 ATGAGCAGTACAGGCACTGGGGCCGAGCCACCAAGCAAGATGCGGAAGCTCT 540
 DB 481 ATGAGCAGTACAGGCACTGGGGCCGAGCCACCAAGCAAGATGCGGAAGCTCT 540
 QY 541 TCAGTTTCACACCAAGCCCTGGAACCTGACCTGCAAGGACTTGTCTCTCCAGGCCCTTAAGGG 600
 DB 541 TCAGTTTCACACCAAGCCCTGGAACCTGACCTGCAAGGACTTGTCTCTCCAGGCCCTTAAGGG 600
 QY 601 AGTCCAGTCTTACCTGTGAGGAGCCTGAGCGGAGCTAGGCTCTCTCCAGCAACAC 660
 DB 601 AGTCCAGTCTTACCTGTGAGGAGCCTGAGCGGAGCTAGGCTCTCTCCAGCAACAC 660
 QY 661 TCCGCTCAGCCCTTGGCAATCCACCAATCATCTGTAATCTGATCTTTTATACACAA 720
 DB 661 TCCGCTCAGCCCTTGGCAATCCACCAATCATCTGTAATCTGATCTTTTATACACAA 720
 QY 721 ATACGAAAAGCCAGCTTGAA 740
 DB 721 ATACGAAAAGCCAGCTTGAA 740

RESULT 2
 ABK87966
 ID ABK87966 standard; cDNA; 740 BP.
 AC ABK87966;
 XX
 DT 07-OCT-2002 (first entry)
 XX
 DE Human caspase recruitment domain-5 (CARD-5) cDNA.
 KW Human; gene; ss; caspase recruitment domain-5; CARD-5; antiinflammatory;
 KW immunosuppressive; caspase; cysteinyl aspartate-specific proteinase;
 KW apoptosis; nuclear factor-kappa B; NF-kappaB; transcription factor;
 KW cell proliferation; gene therapy; immune disorder;
 KW chronic inflammatory disease; Hashimoto's thyroiditis; graft rejection;
 KW sarcoidosis; atopy; asthma; allergy; glomerular nephritis;
 KW human immunodeficiency virus; HIV; bacterial infection; tuberculosis;
 KW lepromatous leprosy; autoimmune disorder; systemic lupus erythematosus;
 KW arthritis; cell depletion; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; spinal muscular atrophy; haematologic disease;
 KW myelodysplastic syndrome; aplastic anaemia; myocardial infarction;
 KW stroke.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 54..641
 FT /tag= a
 FT /product= "CARD-5"
 XX
 PN WO200244354-A2.
 XX
 PD 06-JUN-2002.
 XX
 PF 29-NOV-2001; 2001WO-US44894.
 XX
 PR 01-DEC-2000; 2000US-0728721.
 PR 24-APR-2001; 2001US-0841879.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Bertin J;
 XX
 DR WPI; 2002-557538/59.
 DR P-PSDB; AAU99353.
 XX
 DR Novel isolated murine or human caspase recruitment domain (CARD)-5
 PT polypeptide, useful for treating immune disorders such as Hashimoto's
 PT thyroiditis, graft rejection, allergy, glomerular nephritis,
 PT tuberculosis -
 XX

PS	Claim 7; Fig 3; 10pp; English.
XX	The invention discloses the isolated polypeptides, and encoding nucleic acids, of murine and human caspase recruitment domain (CARD)-5. Caspases (cysteiny aspartate-specific proteinases) are central to the apoptotic program and responsible for the degradation of cellular proteins that lead to the morphological changes seen in cells undergoing apoptosis. CC Caspases interact with other caspases via their CARDS and different CC subtypes of CARDS may confer binding specificity. CARD-5 is an CC intracellular protein that is predicted to be involved in regulating CC caspase activation. CARD-5 activates the nuclear factor-kappa B CC (NF-kappaB) transcription factor pathway and binds the CARDS of CC caspase-1, CARD-7 and itself. CARD-5 can, therefore, modulate CARD-5 CC activity and NF-kappaB activation, regulate cell growth and cell death CC and be used in gene therapy. The CARD-5 polypeptides are useful for CC identifying compounds which bind to them and modulate their activity and CC for detecting the presence of CARD-5 in a sample. CARD-5 polypeptides, CC nucleic acids, antibodies and modulators of CARD-5 expression or activity CC can be used to treat immune disorders such as chronic inflammatory CC diseases and disorders, Hashimoto's thyroiditis, graft rejection, CC sarcoidosis, atopic conditions (such as asthma and allergy), glomerular CC nephritis, human immunodeficiency virus (HIV) and bacterial infections CC (including tuberculosis and lepromatous leprosy) and in screening and CC detection assays. Modulators of CARD-5 activity or expression are also CC useful for treating autoimmune disorders, such as systemic lupus CC erythematosus and arthritis, cell depletion, neurological disorders, CC such as Alzheimer's disease, Parkinson's disease and spinal muscular CC atrophy, haematologic diseases, such as myelodysplastic syndrome and CC aplastic anaemia, myocardial infarction and stroke. The sequence CC presented is the human caspase recruitment domain-5 (CARD-5) cDNA.
XX	
SQ	Sequence 740 BP; 146 A; 238 C; 236 G; 120 T; 0 other;
	Query Match 100.0%; Score 740; DB 24; Length 740;
	Best Local Similarity 100.0%; Pred. No. 1.7e-147;
	Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 CGCGTCGGCTGCAGCGGGTACGGGCGGACGGCCGGGGATCCTGGAGCCATGGGCG 60
Db	1 CGCGTCGGCTGCAGCGGGTGAGCGGCGGACGGCCGGGGATCTCTGGAGCCATGGGCG 60
Qy	61 GCGCGCGCACGCCCATCTGGATGCGTGGAGAACCTCACCGCCGAGGAGGTCAAGAAGT 120
Db	61 GCGCGCGCACGCCCATCTGGATGCGTGGAGAACTTACC CGCGAGAGGTCTCAAGAAGT 120
Qy	121 TCAGAGTGAAAGCTGTGTGGTGCCTGTCGCGAGGGCTTA CGGCGCATATCCCGGGGCG 180
Db	121 TCAGAGTGAAAGCTGTGTGGTGCCTGTCGCGAGGGCTACGGGCGCATCCCGGGGCG 180
Qy	181 CGCTGCTGTCCATGGACGCCCTTGACACTTACCGACAAGCTGGTACGTTCTTACCTGGAGA 240
Db	181 CGCTGCTGTCCATGGACGCCCTTGACACTTACCGACAAGCTGGTACGTTCTTACCTGGAGA 240
Qy	241 CCTACGGCGCGAGCTCACCGCTAACGTGTCTGCGCAATGGGCCCTGCAGAGATGGCGG 300
Db	241 CCTACGGCGCGAGCTCACCGCTAACGTGTCTGCGCAATGGGCCCTGCAGAGATGGCGG 300
Qy	301 GGCAAGTGCAGCGGCGCACGCA CAGGGCTCTGGAGCGCGCGCAGCTGGATTCAGAGCCC 360
Db	301 GGCAAGTGCAGCGGCGCACGCA CAGGGCTCTGGAGCGCGCGCAGCTGGGATTCAGAGCCC 360
Qy	361 CTCCTCAGTCGGCAGCCAAAGCCAGGCGCTGATCTTATAGACACGACCGGGCTGCGCTTA 420
Db	361 CTCCTCAGTCGGCAGCCAAAGCCAGGCGCTGATCTTATAGACACGACCGGGCTGCGCTTA 420
Qy	421 TCGCGAGGGTCAAAA CGTTGAGTGGCTGCTGGATGCTCTGTACGGGAAGGTCTCTGACGG 480
Db	421 TCGCGAGGGTCAAAA CGTTGAGTGGCTGCTGGATGCTCTGTACGGGAAGGTCTCTGACGG 480
Qy	481 ATGAGCAGTACAGGCA GTGCGGCCGAGCCCA CCAACCCAGCAAGATGCGGAAGTCTT 540
Db	481 ATGAGCAGTACAGGCA GTGCGGCCGAGCCCA CCAACCCAGCAAGATGCGGAAGTCTT 540

comprise the ability to induce caspase activation. The methods are useful for treating a disorder associated with inappropriate apoptosis or inappropriate inflammation. The methods are useful for treating disorders associated with an undesirably low rate of apoptosis such as cancer (preferably follicular lymphoma, chronic myelogenous leukaemia, melanoma, colon cancer, lung carcinoma, etc), viral infections, autoimmune diseases caused by low levels of apoptosis (e.g. systemic lupus erythematosus, immune-mediated glomerulonephritis, and arthritis). The methods are also useful for treating disorders with undesirably high rates of apoptosis such as human immunodeficiency virus (HIV) infection, Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis (ALS), retinitis pigmentosa, spinal muscular atrophy, various forms of cerebellar degeneration, anaemia associated with chronic disease, aplastic anaemia, chronic neutropenia, myelodysplastic syndromes, myocardial infarction, stroke, and various inflammatory disorders (e.g. Crohn's disease, reactive arthritis, insulin dependent diabetes mellitus, multiple sclerosis, psoriasis, graft rejection, allergic rhinitis, food allergies, conjunctivitis, glomerular nephritis, etc). The present sequence encoding human CARD-5 is used to identify cDNA encoding CARD-7.

XX Sequence 740 BP; 146 A; 238 C; 236 G; 120 T; 0 other;

Query Match 100.0%; Score 740; DB 25; Length 740;
 Best Local Similarity 100.0%; Pred. No. 1.7e-147;
 Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGTCCGGCTCAGCGGGTGAGCGGGGAGCGCGCGGGATCTCGAGCCATGGGCG 60
 DB 1 CGCGTCCGGCTCAGCGGGTGAGCGGGGAGCGCGGGGATCTCGAGCCATGGGCG 60

QY 61 GCGCGCGGACGCCATCTCTGGATGGCTGGGAACTCTGACCCCGAGGAGCTCAAGAAGT 120
 DB 61 GCGCGCGGACGCCATCTCTGGATGGCTGGGAACTCTGACCCCGAGGAGCTCAAGAAGT 120

QY 121 TCAAGCTGAAGTGTGTGCTGGTCCGCTCGCGAGGGGCTACGGGGCATCCCGGGGGCG 180
 DB 121 TCAAGCTGAAGTGTGTGCTGGTCCGCTCGCGAGGGGCTACGGGGCATCCCGGGGGCG 180

QY 181 CGCTGTGTCATGAGCGCTTGGACCTCACCAGCAAGCTGTCTACCTTGGAGA 240
 DB 181 CGCTGTGTCATGAGCGCTTGGACCTCACCAGCAAGCTGTCTACCTTGGAGA 240

QY 241 CTTAGCGCGCGAGCTCAGCTTACCTGCTCGGCGACATGCGGCTCGAGGAGTGGCGG 300
 DB 241 CTTAGCGCGCGAGCTCAGCTTACCTGCTCGGCGACATGCGGCTCGAGGAGTGGCGG 300

QY 301 GGCAGCTGAGCGGCGCAACGAGGCTCTGGAGCGCGCGAGTGGGATCCAGGCGC 360
 DB 301 GGCAGCTGAGCGGCGCAACGAGGCTCTGGAGCGCGCGAGTGGGATCCAGGCGC 360

QY 361 CTCCTCAGTCCGAGCGAGCGCGCTGCTTATAGACCGACCGGGCTGGCTTA 420
 DB 361 CTCCTCAGTCCGAGCGAGCGCGCTGCTTATAGACCGACCGGGCTGGCTTA 420

QY 421 TCGCGAGGCTCAAAACGTTGAGTGTCTGTGATGCTCTGTACGGGAAGTCTTCAGCG 480
 DB 421 TCGCGAGGCTCAAAACGTTGAGTGTCTGTGATGCTCTGTACGGGAAGTCTTCAGCG 480

QY 481 ATGAGCAGTACCAAGGAGTCCGGGCGGAGCCCAACCAACCAAGAGATGCGGAAGCTCT 540
 DB 481 ATGAGCAGTACCAAGGAGTCCGGGCGGAGCCCAACCAACCAAGAGATGCGGAAGCTCT 540

QY 541 TCAGTTTCAACAGCTGGAATGCACTGCAAGGAGTGTCTTCCAGGCGCTTAAGG 600
 DB 541 TCAGTTTCAACAGCTGGAATGCACTGCAAGGAGTGTCTTCCAGGCGCTTAAGG 600

QY 601 AGTCCAGTCTTACCTGGTGGAGGCTCGAGCGGAGCTGAGCTCTTCCAGCAAC 660
 DB 601 AGTCCAGTCTTACCTGGTGGAGGCTCGAGCGGAGCTGAGCTCTTCCAGCAAC 660

QY 661 TCGGCTCAGCGCGCTGCAATCCCAACCAATCATCTGAATCTGATCTTTTATACAAAT 720
 DB 661 TCGGCTCAGCGCGCTGCAATCCCAACCAATCATCTGAATCTGATCTTTTATACAAAT 720

Db 661 TCGGCTCAGCGCGCTGCAATCCCAACCAATCATCTGAATCTGATCTTTTATACAAAT 720

QY 721 ATACGAAAAGCCAGCTTGAA 740
 ||||||||||||||||||
 Db 721 ATACGAAAAGCCAGCTTGAA 740
 ||||||||||||||||||

RESULT 5
 AAD03890
 ID AAD03890 standard; cDNA; 770 BP.
 AC AAD03890;
 XX 02-JUL-2001 (first entry)
 DT Human target of methylation-induced silencing-1 (TMS1) cDNA.
 DE Human; target of methylation-induced silencing-1; TMS1; cytostatic;
 KW antiproliferative; apoptosis inducer; gene therapy; CpG island;
 KW caspase-recruiting domain; CARD; cancer; breast; ss.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 75..662
 FT /*tag= a
 FT /product= "Human TMS1 protein"

XX WO200129235-A2.
 PD 26-APR-2001.
 XX 18-OCT-2000; 2000WO-US28747.
 PF 18-OCT-1999; 99US-0159975.
 PR (UYEM-) UNIV EMORY.
 XX Vertino PW;
 DR WPI; 2001-290922/30.
 DR P-PSDB; AAE00588.
 XX Novel gene TMS1, transcriptionally silenced due to increased methylation useful for identifying subject at risk of developing tumor characterized by abnormal methylation, for treating cancer by inducing apoptosis

XX Claim 68; Page 113-114; 124pp; English.
 CC The invention relates to identification of target of methylation-induced silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to abnormal methylation of a CpG island in its 5' regulatory region. TMS1 consists of a carboxy terminal caspase-recruiting domain (CARD) and plays a role in induction of apoptosis. TMS1 gene and protein are useful as tools for diagnosing and treating a subject at risk of developing cancer (e.g. breast cancer) characterised by abnormal CpG methylation or abnormally low levels of TMS1 expression products. Unique fragments of TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.
 CC TMS1 molecule is also useful for treating abnormal cell proliferation by increasing TMS1 polypeptide level to an above normal level. The CpG island region of TMS1 or its fragments are used to study the methylation patterns apart from any coding region contained in it.
 CC The present sequence is human target of methylation-induced silencing-1 (TMS1) cDNA.

XX Sequence 770 BP; 157 A; 247 C; 241 G; 125 T; 0 other;

Query Match 99.2%; Score 734; DB 22; Length 770;
 Best Local Similarity 100.0%; Pred. No. 3.2e-146;
 Matches 734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 CGGCTGCAGCGGGGTGAGCGCGGCGGCGGCGGATCCTGGAGCCATGGGCGCGCGC 66

Db 28 CGGCTGACGGGGTGGAGCGGCGAGCGCGGGGATCTCTGGAGCATGGGGCGCGCG 87
Qy 67 GCGACGCCATCTGGATGGCTGGAGAACTTGACCGCGAGGAGCTCAAGAGTTCAAGC 126
Db 88 GCGACGCCATCTGGATGGCTGGAGAACTTGACCGCGAGGAGCTCAAGAGTTCAAGC 147
Qy 127 TGAAGCTGCTGCTGGTGGCTGGCGAGGGCTACGGGGCGCATCCCGGGGGCGCTGC 186
Db 148 TGAAGCTGCTGCTGGTGGCTGGCGAGGGCTACGGGGCGCATCCCGGGGGCGCTGC 207
Qy 187 TGTTCATGAGCGCTTTGGACCTCACCGACAAGCTGTGCTGAGCTTCTACCTGGAGACCTAG 246
Db 208 TGTTCATGAGCGCTTTGGACCTCACCGACAAGCTGTGCTGAGCTTCTACCTGGAGACCTAG 267
Qy 247 GCGCCGAGCTACCGCTACGCTGCTGGCGGACATGGGGCTGACGAGATGGCGGGCGAGC 306
Db 268 GCGCCGAGCTACCGCTACGCTGCTGGCGGACATGGGGCTGACGAGATGGCGGGCGAGC 327
Qy 307 TGCAGCGCGCCAGCACCGAGGCTCTGGAGCGCGCGAGCTGGATCCAGGCGCCCTCCTC 366
Db 328 TGCAGCGCGCCAGCACCGAGGCTCTGGAGCGCGCGAGCTGGATCCAGGCGCCCTCCTC 387
Qy 367 AGTCGCGAGCGCAAGCGAGCGCTGCACTTTATAGACGACCGGGCTGGCTTATCGCGA 426
Db 388 AGTCGCGAGCGCAAGCGAGCGCTGCACTTTATAGACGACCGGGCTGGCTTATCGCGA 447
Qy 427 GGGTCACAAACGTTGAGTGGCTGGATGCTCTGTACGGAAAGTCTTGACGAGTACG 486
Db 448 GGGTCACAAACGTTGAGTGGCTGGATGCTCTGTACGGAAAGTCTTGACGAGTACG 507
Qy 487 AGTACCAGCAGTGGCGGGCGGAGCCACCAACCAAGCAAGATGCGGAAGCTCTTCAGTT 546
Db 508 AGTACCAGCAGTGGCGGGCGGAGCCACCAACCAAGCAAGATGCGGAAGCTCTTCAGTT 567
Qy 547 TCACACAGCCTGGAACTGGAACTGGAACTGGCTCTCCAGGAGCTTCTCCAGGAGTCCC 606
Db 568 TCACACAGCCTGGAACTGGAACTGGAACTGGCTCTCCAGGAGCTTCTCCAGGAGTCCC 627
Qy 607 AGTCTACCTGGTGGAGGAGCTGGAGCGAGCTGAGGCTCTTCCAGCAACACTCCCGT 666
Db 628 AGTCTACCTGGTGGAGGAGCTGGAGCGAGCTGAGGCTCTTCCAGCAACACTCCCGT 687
Qy 667 CAGCCCTCGCAATCCCAATATCATCTGAACTGATCTTTTATACAAATATACGA 726
Db 688 CAGCCCTCGCAATCCCAATATCATCTGAACTGATCTTTTATACAAATATACGA 747
Qy 727 AAAGCCAGCTTGAA 740
Db 748 AAAGCCAGCTTGAA 761

RESULT 6
ID AAL47126
XX AAL47126 standard; DNA; 782 BP.
XX
AC AAL47126;
XX
DT 20-AUG-2002 (first entry)
XX
DE Pyrin domain containing protein Pycard coding sequence.
XX
KW Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
KW antiarteriosclerotic; antiporotic; antibacterial; virucide;
KW neuroprotective; antarthritic; antirheumatic; antiasthmatic;
KW nephrotropic; osteoparitic; nootropic; intracellular signal transduction;
KW inflammation; Alzheimer's disease; infection; psoriasis; asthma;
KW arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
KW osteoarthritis; glomerulonephritis; gene; ds.
XX
OS Unidentified.
XX
PN W0200240668-A2.

XX 23-MAY-2002.
XX 30-OCT-2001; 2001WO-EP12545.
XX 15-NOV-2000; 2000DE-1056687.
XX 30-NOV-2000; 2000DE-1059595.
XX (APOT-) APOTECH RES & DEV LTD.
XX
XX Tschoopp J, Martinon F;
XX WPI; 2002-427093/45.
XX P-PSDB; AAO17854.
XX
XX New DNA encoding protein with pyrin domain, useful for treating
XX diseases involving impaired signal transduction, particularly
XX inflammation, also proteins and antibodies -
XX
XX Claim 5; Fig 1; 116pp; German.
XX
XX The present invention relates the DNA and their encoded proteins, where
XX the proteins contain at least one PYD (pyrin) domain. These can be used
XX to treat diseases associated with impaired intracellular signal
XX transduction, particularly inflammation such as psoriasis,
XX arteriosclerosis, bacterial or viral infections (particularly meningitis
XX and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
XX sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
XX and Parkinson's diseases. The present sequence is a coding sequence of
XX the invention.
XX
XX Sequence 782 BP; 158 A; 251 C; 246 G; 127 T; 0 other;
XX
Query Match 99.2%; Score 734; DB 24; Length 782;
Best Local Similarity 100.0%; Pred. No. 3.2e-146;
Matches 734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 7 CGGCTGACGGGGTGGAGCGGCGAGCGCGGGATCTTGAGCCATGGGGCGCGCG 66
Db 40 CGGCTGACGGGGTGGAGCGGCGAGCGCGGGATCTTGAGCCATGGGGCGCGCG 99
Qy 67 GCGAGCCCATCTGGATGCTGGAGAACTTGACCGCGAGGAGCTCAAGAGTTCAAGC 126
Db 100 GCGAGCCCATCTGGATGCTGGAGAACTTGACCGCGAGGAGCTCAAGAGTTCAAGC 159
Qy 127 TGAAGCTGCTGCTGGTGGCTGGAGGAGCTACGGCGGAGTCCCGGGGGCGCTGC 186
Db 160 TGAAGCTGCTGCTGGTGGCTGGAGGAGCTACGGCGGAGTCCCGGGGGCGCTGC 219
Qy 187 TGTTCATGAGCGCTTTGGACCTCACCGACAAGCTTGTGCTGAGCTTCTACCTGGAGACCTAG 246
Db 220 TGTTCATGAGCGCTTTGGACCTCACCGACAAGCTTGTGCTGAGCTTCTACCTGGAGACCTAG 279
Qy 247 GCGCGAGCTCACCGCTAACGCTGCTGCGGACATGGGCTTCAGAGATGGCGGGCGAGC 306
Db 280 GCGCGAGCTCACCGCTAACGCTGCTGCGGACATGGGCTTCAGAGATGGCGGGCGAGC 339
Qy 307 TGCAGCGCGCACCGCACCGAGGCTTGGAGCGCGCGAGCTGGGATCCAGGCCCCCTCCTC 366
Db 340 TGCAGCGCGCACCGCACCGAGGCTTGGAGCGCGCGAGCTGGGATCCAGGCCCCCTCCTC 399
Qy 367 AGTCGCGAGCGCAAGCGAGGCTGCACTTTATAGACGACCGGGCTCGCTTATCGCGA 426
Db 400 AGTCGCGAGCGCAAGCGAGGCTGCACTTTATAGACGACCGGGCTCGCTTATCGCGA 459
Qy 427 GGGTCACAAACGTTGAGTGGCTGGATGCTCTGTACGGAAAGTCTTGACGAGTACG 486
Db 460 GGGTCACAAACGTTGAGTGGCTGGATGCTCTGTACGGAAAGTCTTGACGAGTACG 519
Qy 487 AGTACCAGGAGTGGCGGGCGAGCCCAACCCAGCAAGATGCGGAAGCTTCTTCAGTT 546
Db 520 AGTACCAGGAGTGGCGGGCGAGCCCAACCCAGCAAGATGCGGAAGCTTCTTCAGTT 579

547	TCACACCGCCTGGAACTCGAGCCTCGAAGGACTTGTCTCTCCAGGCCCTTAAGGAGTCCC	606
Qy		
580	TCACACCGCCTGGAACTCGAGCCTCGAAGGACTTGTCTCTCCAGGCCCTTAAGGAGTCCC	639
Db		
607	AGTCCTACCTGGTGGAGGACCTGGAGCGGAGCTGAGGCTCCTCCAGCAACACTCCGGT	666
Qy		
640	AGTCCTACCTGGTGGAGGACCTGGAGCGGAGCTGAGGCTCCTCCAGCAACACTCCGGT	699
Db		
667	CAGCCCTGGCAATCCCAACAAATCATCTCGAATCTGATCTTTTATACAAATATACGA	726
Qy		
700	CAGCCCTGGCAATCCCAACAAATCATCTCGAATCTGATCTTTTATACAAATATACGA	759
Db		
727	AAAGCCAGCTTGAA	740
Qy		
760	AAAGCCAGCTTGAA	773
Db		

RESULT 7

AAAS59817	
ID	AAAS59817 standard; cDNA; 745 BP.
XX	
XX	AAAS59817;
XX	
DT	16-JAN-2002 (first entry)
XX	
DE	Human novel cytokine encoding cDNA 790CIP2B_1 #1.
XX	
XX	Human; ss; cytokine; cell proliferation; cell differentiation;
KW	antiinflammatory; stem cell growth factor; activin; inhibin; cancer;
KW	nervous system disease; neuropathy; Alzheimer's disease;
KW	Parkinson's disease; Huntington's disease; spinal cord disorder;
KW	head trauma; stroke; myeloid cell disorder; lymphoid cell disorder;
KW	platelet disorder; thrombocytopaenia; stem cell disorder;
KW	aplastic anaemia; tissue regeneration; wound healing; ulcer;
KW	osteoporosis; osteoarthritis; bone degenerative disorder;
KW	periodontal disease; fibrosis; reperfusion; immune disorder; SCID;
KW	severe combined immunodeficiency; infection; autoimmune disorder;
KW	multiple sclerosis; rheumatoid arthritis; diabetes mellitus; allergy;
KW	asthma; coagulation disorder; haemophilia; sepsis; nephritis;
KW	inflammatory bowel disease; food supplement.
XX	
OS	Homo sapiens.
XX	
PN	WQ200175093-A1.
XX	
XX	11-OCT-2001.
PD	
XX	30-MAR-2001; 2001WO-US10484.
XX	
PR	31-MAR-2000; 2000US-0540217.
PR	23-AUG-2000; 2000US-0649167.
PR	22-SEP-2000; 2000US-0668680.
PR	23-OCT-2000; 2000US-0695618.
PR	30-NOV-2000; 2000US-0728711.
PR	14-MAR-2001; 2000US-0728711.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J, Xu C;
PI	Yang Y, Zaho QA, Chen R, Wang D, Goodrich RW, Liu C, Drmanac RT;
XX	
XX	WPI: 2001-626432/72.
DR	P-PSDB; AAU68525.
DR	
XX	
PT	New polypeptides and nucleic acids, useful for diagnosis, treatment of
PT	inflammatory, autoimmune, neurological, myeloid or lymphoid cell, bone
PT	degenerative disorders, cancer and promoting wound healing -
XX	
PS	Claim 1; Page 150-151; 336pp; English.
XX	
CC	The invention relates to isolated human polypeptides (which may be
CC	cytokines) and the polynucleotides encoding them. The protein is useful
CC	for identifying a compound which binds to it (e.g. modulators, agonists

PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 useful for preventing, diagnosing and/or treating colorectal cancers -
 Claim 1; Page 2955; 9803pp; English.

CC AAH32943 to AAH37195 and AG773514 to AAG77788 represent human colon
 cancer-associated nucleic acid molecules (N) and proteins (P), where
 the proteins are collectively known as colon cancer antigens. The colon
 cancer antigens have cytosolic activity and can be used in gene
 therapy and vaccine production. N and P may be used in the prevention,
 diagnosis and treatment of diseases associated with inappropriate P
 expression. For example, N and P may be used to treat disorders
 associated with decreased expression by rectifying mutations or deletions
 in a patient's genome that affect the activity of P by expressing
 inactive proteins or to supplement the patients own production of P.
 Additionally, N may be used to produce the colon cancer-associated Ps,
 by inserting the nucleic acids into a host cell and culturing the cell
 to express the proteins. N and P can be used in the prevention, diagnosis
 and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 and AAB77789 represent sequences used in the exemplification of the
 present invention.

CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.

XX Sequence 806 BP; 163 A; 253 C; 260 G; 128 T; 2 other;

Query Match 97.5%; Score 721.2; DB 23; Length 806;
 Best Local Similarity 99.6%; Pred. No. 1.7e-143; Indels 1; Gaps 1;
 Matches 731; Conservative 2; Mismatches 0;

7 CGGCTGCACGCGGGTGCAGCGCGCGGCGGATCCTGCAGCCATGGCGCGCGCG 66
 68 CGGCTGCAGCGGGTGCAGCGCGCGGCGGATCCTGCAGCCATGGCGCGCGCG 127
 67 GCGAGCCCATCTGCATGCGTGGAGAACCTGACCGCGGAGAGCTCAAGAGTTCAAGC 126
 128 GCGAGCCCATCTGCATGCGTGGAGAACCTGACCGCGGAGAGCTCAAGAGTTCAAGC 187
 127 TGAAGTGTGTCGTCGCGTGGCGGCGGAGGCTACGGGCGCATCCGCGGGCGCGTGC 186
 188 TGAAGTGTGTCGTCGCGTGGCGGCGGAGGCTACGGGCGCATCCGCGGGCGCGTGC 247
 187 TGTTCATGACGCGCTTGCACCTCACCAGCAAGCTGTGTCAGCTTCTACCTGGAGACCTAG 246
 248 TGTTCATGACGCGCTTGCACCTCACCAGCAAGCTGTGTCAGCTTCTACCTGGAGACCTAG 307
 247 GCGCGAGCTACCGCTAACGTCGTGCGGACATGGGCTGCAGGAGATGGCGGGCAGC 306
 308 GCGCGAGCTACCGCTAACGTCGTGCGGACATGGGCTGCAGGAGATGGCGGGCAGC 367
 307 TGCAGGCGCCACGACACCGAGGCTGTGAGCGCGGCGGATCCAGGCGCCCTCCTC 366
 368 TGCAGGCGCCACGACACCGAGGCTGTGAGCGCGGCGGATCCAGGCGCCCTCCTC 427
 367 AGTTCGCGCAGCAGCAGCGCTGCATCTTATAGACAGCAGCGGCTGCCTTATCGCA 426
 428 AGTTCGCGCAGCAGCAGCGCTGCATCTTATAGACAGCAGCGGCTGCCTTATCGCA 487
 427 GGGTCAAAAGCTTGTAGTGGTGTGATGCTGTCTGTACGGGAAGTTCCTGACGGATGAGC 486
 488 GGGTCAAAAGCTTGTAGTGGTGTGATGCTGTCTGTACGGGAAGTTCCTGACGGATGAGC 547
 487 AGTACGAGCAGTGGCGGCGGACCCACCAAGCAAGATGCGGAGCTTTCAGTT 546
 548 AGTACGAGCAGTGGCGGCGGACCCACCAAGCAAGATGCGGAGCTTTCAGTT 607
 547 TCACACGAGCTGGAAGTGAAGTCTCAAGGAGTTCCTTCCAGGCGCTTTCAGTT 606
 608 TCACACGAGCTGGAAGTGAAGTCTCAAGGAGTTCCTTCCAGGCGCTTTCAGTT 667
 607 AGTTCCTACCTGTTGGAGGAGCTTGGAGCGGAGTGGGCTCCTTCCAGCAACACCTCCGT 666

Db 668 AGTCTACTCTGGTGAGGACCTKGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGGT 727
 Qy 667 CAGCCCTTGGCAATCCCAACCAATCATCTGAATCTGATCTTTTATACAAATACGA 726
 Db 728 CA-CCCTTGGCAATCCCAACCAATCATCTGAATCTGATCTTTTATACAAATACGA 786
 Qy 727 AAAGCCAGCTTGAA 740
 Db 787 AAAGCCAGCTTGAA 800

RESULT 10
 AAC77884
 ID AAC77884 standard; cDNA; 811 BP.
 XX
 AC AAC77884;
 XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human cancer associated gene sequence SEQ ID NO:278.
 XX
 KW Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
 KW antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
 KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
 KW dermatologic; neuroprotective; thrombolytic; coagulant; nootropic;
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening; ss.
 XX
 OS Homo sapiens.
 XX
 FN WO2000055350-A1.
 XX
 PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000WO-US05882.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2000-587533/55.
 DR P-PSDB; AAB43675.
 PT Novel isolated nucleic acids comprising sequences encoding peptides
 useful for treating or diagnosing e.g. cancer -
 XX
 PS Claim 1; Page 841; 2352pp; English.
 XX
 CC AAC77607 to AAC78448 encode the human cancer associated proteins given
 in AAB43398 to AAB44239. The proteins can have activities based on the
 tissues and cells the genes are expressed in. Example of activities
 include: cytostatic; proliferative; vulnery; immunomodulator;
 CC antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
 CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
 CC dermatologic; neuroprotective; cardiant; thrombolytic; coagulant;
 CC nootropic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to

CC	AAC78457 and AAB44240 represent sequences used in the exemplification of the present invention.
XX	
XX	
SQ	Sequence 811 BP; 170 A; 253 C; 259 G; 128 T; 1 other;
	Query Match 94.3%; Score 697.6; DB 21; Length 811;
	Best Local Similarity 99.5%; Pred. No. 1.6e-138;
	Matches 730; Conservative 1; Mismatches 0; Indels 3; Gaps 3;
Qy	7 CGGCTCAGCGGGTGAGCGGGCGGAGCGCGGGATCCTGAGGCATGGGGCGCGCGC 66
Db	68 CGGCTCAGCGGGTGAGCGGGCGGAGCGCGGGATCCTGAGGCATGGGGCGCGCGC 127
Qy	67 GGCACGCCATCTGGATGGCGTGGAGAACTGACCGCGGAGGAGCTCAAGAAAGTTCAAGC 126
Db	128 GGCACGCCATCTGGATGGCGTGGAGAACTGACCGCGGAGGAGCTCAAGAAAGTTCAAGC 187
Qy	127 TGAAGCTGTGTGGTGGCGTGGCGGAGGGCTACGGGCGCATGCCGGGGCGCGCTGC 186
Db	188 TGAAGCTGTGTGGTGGCGTGGCGGAGGGCTACGGGCGCATGCCGGGGCGCGCTGC 247
Qy	187 TGTCCATGACCGCTTGGACCTCACGCAAGCTGGTCACTTACCTGGAGACCTAGC 246
Db	248 TGTCCATGACCGCTTGGACCTCACGCAAGCTGGTCACTTACCTGGAGACCTAGC 307
Qy	247 GGGCCGAGCTCACCGCTAACGCTGCGCGACATGGGCGCTGCAGGAGATGGCGGGCAGC 306
Db	308 GGGCCGAGCTCACCGCTAACGCTGCGCGACATGGGCGCTGCAGGAGATGGCGGGCAGC 367
Qy	307 TGCAGGCGCCACGACACCGAGGCTCTGGAGCGCGCGCATCTGGATGCCAGGCCCTCTCTC 366
Db	368 TGCAGGCGCCACGACACCGAGGCTCTGGAGCGCGCGCATCTGGATGCCAGGCCCTCTCTC 427
Qy	367 AGTCGGCAGCCAGCCAGCGCTGCATTTATAGACCAGCAGCGGCTGGCTTATCGCA 426
Db	428 AGTCGGCAGCCAGCCAGCGCTGCATTTATAGACCAGCAGCGGCTGGCTTATCGCA 487
Qy	427 GGGTCACAAACGTTGAGTGGCTGCTGGATGCTCTGTACGGGAAGTCTTACGGATGAGC 486
Db	488 GGGTCACAAACGTTGAGTGGCTGCTGGATGCTCTGTACGGGAAGTCTTACGGATGAGC 547
Qy	487 AGTACAGGAGTGGCGGCGGAGCCACCAACCCAGCAAGATGCGGAAGCTCTTCAAGTT 546
Db	548 AGTACAGGAGTGGCGGCGGAGCCACCAACCCAGCAAGATGCGGAAGCTCTTCAAGTT 606
Qy	547 TCACACAGCGTGAAGTGAAGCTGCAAGACTTCTCTCCAGGCCCTTAAGGAGTCCC 606
Db	607 TCACACAGCGTGAAGTGAAGCTGCAAGACTTCTCTCCAGGCCCTTAAGGAGTCCC 666
Qy	607 AGTCCTTACCTGGTGGAGGACCTGGAGCGAGCTGAGGCTCCTTCCAGCAACACTCCGGT 666
Db	667 AGTCCTTACCTGGTGGAGGACCTGGAGC-GAGCTGAGGCTCCTTCCAGCAACACTCCGGT 725
Qy	667 CAGCCCTCGCAATCCCAACCAATCATCTGATCTGATCTTTTATACAAATATACGA 726
Db	726 CA-SCCTCGCAATCCCAACCAATCATCTGATCTGATCTTTTATACAAATATACGA 784
Qy	727 AAAGCCAGCTTGA 740
Db	785 AAAGCCAGCTTGA 798
RESULT 11	
AAD03906	
ID	AAD03906 standard; cDNA; 713 BP.
XX	
AC	AAD03906;
XX	
DT	02-JUL-2001 (first entry)
XX	
DE	Alternatively spliced form of human TMS1 cDNA (lacking exon2).
XX	
KW	Human; target of methylation-induced silencing-1; TMS1; cytosstatic;

KW	antiproliferative; apoptosis inducer; gene therapy; CpG island;
KW	caspace-recruiting domain; CARD; cancer; breast; ss.
XX	
OS	Homo sapiens.
XX	
PH	Key Location/Qualifiers
FT	75..605
FT	/tag= a
FT	/product= "Alternatively spliced form of human TMS1
FT	protein lacking exon2"
XX	
PN	WO200129235-A2.
XX	
PD	26-APR-2001.
XX	
PF	18-OCT-2000; 2000WO-US28747.
XX	
PR	18-OCT-1999; 99US-0159975.
XX	
PA	(UYEM-) UNIV EMORY.
XX	
PI	Vertino PM;
XX	
DR	WPI; 2001-290922/30.
DR	P-PSDB; AAE00594.
XX	
PT	Novel gene TMS1, transcriptionally silenced due to increased
PT	methylation useful for identifying subject at risk of developing tumor
PT	characterized by abnormal methylation, for treating cancer by inducing
PT	apoptosis -
XX	
PS	Claim 68; Page 122; 124pp; English.
XX	
CC	The invention relates to identification of target of methylation-induced
CC	silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to
CC	abnormal methylation of a CpG island in its 5' regulatory region. TMS1
CC	consists of a carboxy terminal caspase-recruiting domain (CARD) and
CC	plays a role in induction of apoptosis. TMS1 gene and protein are useful
CC	as tools for diagnosing and treating a subject at risk of developing
CC	cancer (e.g. breast cancer) characterized by abnormal CpG methylation or
CC	abnormally low levels of TMS1 expression products. Unique fragments of
CC	TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.
CC	TMS1 molecule is also useful for treating abnormal cell proliferation by
CC	increasing TMS1 polypeptide level to an above normal level. The CpG
CC	island region of TMS1 or its fragments are used to study the methylation
CC	patterns apart from any coding region contained in it.
CC	The present sequence is alternatively spliced form of human target of
CC	methylation-induced silencing-1 (TMS1) cDNA lacking exon2.
XX	
SQ	Sequence 713 BP; 148 A; 224 C; 223 G; 118 T; 0 other;
	Query Match 82.4%; Score 610; DB 22; Length 713;
	Best Local Similarity 92.2%; Pred. No. 5.6e-120;
	Matches 677; Conservative 0; Mismatches 0; Indels 57; Gaps 1;
Qy	7 CGGCTGACGCGGGTGAGCGCGGAGCGCGCGGATCCTGGAGCCATGGGGCGCGCGC 66
Db	28 CGGCTGACGCGGGTGAGCGCGGAGCGCGCGGATCCTGGAGCCATGGGGCGCGCGC 87
Qy	67 GCGAGCCCATCTGGATCGCTGGAGAACTTGAACCTGACCGCGAGAGCTCAAGAGTTCAAGC 126
Db	88 GCGAGCCCATCTGGATCGCTGGAGAACTTGAACCTGACCGCGAGAGCTCAAGAGTTCAAGC 147
Qy	127 TGAAGCTGTGTGGTGGCGTGGAGGGCTACGGGCGCATCCCGGGCGCGCTGC 186
Db	148 TGAAGCTGTGTGGTGGCGTGGAGGGCTACGGGCGCATCCCGGGCGCGCTGC 207
Qy	187 TGTCCATGAGCGCTTGGACCTCACCGCAAGAGCTTGTACCTTGAGAGCTTACG 246
Db	208 TGTCCATGAGCGCTTGGACCTCACCGCAAGAGCTTGTACCTTGTACCTTGAGAGCTTACG 267
Qy	247 GCGCCGAGCTCACCGCTTAACCTGTGCGCGCATATGGGCTTCAGAGATGGCCGGCGAGC 306

Db 268 GGGCCGAGCTACCGCTAACGTGCTGCGCGAATGGGCCCTGACGAGATGGCCGGCGAGC 327
Qy 307 TGCAGCGGCCACGACACAGGGCTCTGGAGCGCGCCAGCTGGATCCAGGCCCCCTCCTC 366
Db 328 TGCAGCGGCCACGACCGA----- 347
Qy 367 AGTCGGCAGCCAAAGCCAGCCCTGCACTTTATATAGACAGACCGGGCTGGCTTATCGCGA 426
Db 348 -----GGCCTGCACCTTTATATAGACAGACCGGGCTGGCTTATCGCGA 390
Qy 427 GGGTCACAAAAGCTGAGTGGCTGGATGCTCTGTACGGGAAGTCTCTGACGGATGAGC 486
Db 391 GGGTCACAAAAGCTGAGTGGCTGGATGCTCTGTACGGGAAGTCTCTGACGGATGAGC 450
Qy 487 AGTACAGGCACTGGCGGCGGAGCCACCAACCAAGCAAGATGGCGAAGCTCTTCAGTT 546
Db 451 AGTACAGGCACTGGCGGCGGAGCCACCAACCAAGCAAGATGGCGAAGCTCTTCAGTT 510
Qy 547 TCACACAGCCTGGAACTGGACCTGCAAGCACTTGTCTCCAGGCGCTTAAGGGAGTCCC 606
Db 511 TCACACAGCCTGGAACTGGACCTGCAAGCACTTGTCTCCAGGCGCTTAAGGGAGTCCC 570
Qy 607 AGTCTACTGCTGGAGGACCTGGAGCGGAGCTGAGGCTCTCTCCAGCAACACTCCGGT 566
Db 571 AGTCTACTGCTGGAGGACCTGGAGCGGAGCTGAGGCTCTCTCCAGCAACACTCCGGT 630
Qy 667 CAGCCCTGGCAATCCCAAAATCATCTGAATCTGATCTTTTATACAAATATACGA 726
Db 631 CAGCCCTGGCAATCCCAAAATCATCTGAATCTGATCTTTTATACAAATATACGA 690
Qy 727 AAAGCCAGCTTGAA 740
Db 691 AAAGCCAGCTTGAA 704

RESULT 12

AAH34053
ID AAH34053 standard; cDNA; 639 BP.
XX
AC AAH34053;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen encoding cDNA SEQ ID NO:1135.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; ss.
XX
OS Homo sapiens.
XX
PN WO200122920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26524.
XX
PR 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
DR WPI; 2001-235357/24.
DR P-PSDB; AAG74648.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX useful for preventing, diagnosing and/or treating colorectal cancers -
XX
PS Claim 1; Page 2955; 9803pp; English.
XX
CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where

CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB7789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.

Qy Sequence 639 BP; 146 A; 198 C; 189 G; 106 T; 0 other;
Query Match 71.8%; Score 531; DB 22; Length 639;
Best Local Similarity 94.9%; Pred. No. 2.9e-103;
Matches 595; Conservative 0; Mismatches 20; Indels 12; Gaps 4;
Qy 124 AGCTGAAGCTGCTGCGTGGCGGCTGCGGAGGCTACGGGCGCATCCCGG-----G 176
Db 2 AGCTGAAGCTGCTGCGTGGCGGCTGCGGAGGCTACGGGCGCGCGACCCATCCTG 61
Qy 177 GCGCGCTGCTGCTCCATGAGCGCTTGACCTCACCAG---CAAGCTGTGAGTCTAC 233
Db 62 GATGCGCTGGAGACCTGACCGCGAGGAGCTCAAGAGTTCAAGTGTGAGTCTAC 121
Qy 234 CTGGAGACCTACGGCGCGAGCTCAACGCTTACGCTGCGGACATGGGCTCGAGGAG 293
Db 122 CTGGAGACCTACGGCGCGAGCTCAACGCTTACGCTGCGGACATGGGCTCGAGGAG 181
Qy 294 ATGGCGGGGAGCTGCGAGCGGCCACGACACAGGCTCTGGAGCGCGCCAGCTGGGATC 353
Db 182 ATGGCGGGGAGCTGCGAGCGGCCACGACACAGGCTCTGGAGCGCGCCAGCTGGGATC 241
Qy 354 CAGGCGCTCTCAGTCGGCGAGCCAGCGGCTGCACTTTATAGACCAGCACCGGGCT 413
Db 242 CAGGCGCTCTCAGTCGGCGAGCCAGCGGCTGCACTTTATAGACCAGCACCGGGCT 301
Qy 414 GGGCTTATCGGAGGCTCAAAAAGCTGAGTGGCTGCTGCTGTACGGGAGGTC 473
Db 302 GGGCTTATCGGAGGCTCAAAAAGCTGAGTGGCTGCTGCTGTACGGGAGGTC 361
Qy 474 CTGACGGATGAGCAGTACAGGCGGCGGCGGAGCCACCAACCAAGCAAGATGGCG 533
Db 362 CTGACGGATGAGCAGTACAGGCGGCGGAGCCAGCGGCTGCTGTCTCCAGGCG 420
Qy 534 AAGCTCTTCAGTTTACACACAGCTGGAACTGGAACTGGAACTGGAACTGGAACTGGAA 593
Db 421 AAGCTCTTCAGTTTACACACAGCTGGAACTGGAACTGGAACTGGAACTGGAACTGGAA 480
Qy 594 CTAAGGAGTCCAGTCTCTACCTGCTGGAGGAGCTGGAGGAGCTGGAGGAGCTGGAGGAG 653
Db 481 CTAAGGAGTCCAGTCTCTACCTGCTGGAGGAGCTGGAGGAGCTGGAGGAGCTGGAGGAG 540
Qy 654 GCAACACTCCGCTCAGCCCCCTGGCAATCCCAACCAATCATCTGATCTGATCTTTTAT 713
Db 541 GCAACACTCCGCTCAG---CCCTGGCAATCCCAACCAATCATCTGATCTGATCTTTTAT 599
Qy 714 ACACAATATACAAAAGCCAGCTTGAA 740
Db 600 ACACAATATACAAAAGCCAGCTTGAA 626

RESULT 13

AAH98638
ID AAC98638 standard; cDNA; 432 BP.
XX

CC abnormally low levels of TMS1 expression products. Unique fragments of
CC TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.
CC TMS1 molecule is also useful for treating abnormal cell proliferation by
CC increasing TMS1 polypeptide level to an above normal level. The CpG
CC island region of TMS1 or its fragments are used to study the methylation
CC patterns apart from any coding region contained in it.
CC The present sequence is alternatively spliced form of human target of
CC methylation-induced silencing-1 (TMS1) cDNA lacking exon3.
CC Note: The present sequence designated as SEQ ID NO:26, is referred
CC as DNA throughout the specification, however this sequence has been
CC referred as protein in claim 106.
XX
SQ Sequence 405 BP; 63 A; 139 C; 147 G; 56 T; 0 other;

Query Match 51.1%; Score 378; DB 22; Length 405;
Best Local Similarity 100.0%; Pred. No. 6.6e-71;
Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 CGGCTGACGGGGTGGAGCGCGGAGCGCGGGGATCCTGGAGCCATGGGGCGCGGC 66
Db 28 CGGCTGACGGGGTGGAGCGCGGAGCGCGGGGATCCTGGAGCCATGGGGCGCGGC 87
QY 67 GCGAGCCCATCTGGATGCGCTGGAGAACCTGACCGCGGAGGCTCAAGAGTTCAAGC 126
Db 88 GCGAGCCCATCTGGATGCGCTGGAGAACCTGACCGCGGAGGCTCAAGAGTTCAAGC 147
QY 127 TGAAGCTGCTGCGGTGCGCTGCGGAGGCTTACGGGCGCATCCCGCGGGCGCGCTGC 186
Db 148 TGAAGCTGCTGCGGTGCGCTGCGGAGGCTTACGGGCGCATCCCGCGGGCGCGCTGC 207
QY 187 TGTCCATGACGCTTGGACCTCACGACAAGCTGTGCTAGCTTCTACCTGGAGACCTAGC 246
Db 208 TGTCCATGACGCTTGGACCTCACGACAAGCTGTGCTAGCTTCTACCTGGAGACCTAGC 267
QY 247 GCGCCGAGCTACCGCTAACGCTGCTCGCGACATGCGGCTGCGAGATGGCGGGCAGC 306
Db 268 GCGCCGAGCTACCGCTAACGCTGCTCGCGACATGCGGCTGCGAGATGGCGGGCAGC 327
QY 307 TCGAGCGCGCCAGCGACCGAGGCTCTGGAGCGCGGCTGCGGATCCAGGCGCCCTCC 366
Db 328 TCGAGCGCGCCAGCGACCGAGGCTCTGGAGCGCGGCTGCGGATCCAGGCGCCCTCC 387
QY 367 AGTGGCGACCCAGCCAG 384
Db 388 AGTGGCGACCCAGCCAG 405

RESULT 15
AAD03889
ID AAD03889 standard; DNA; 2821 BP.
XX
AC AAD03889;
XX
DT 02-JUL-2001 (first entry)
XX
DE Human target of methylation-induced silencing-1 (TMS1) genomic DNA.
XX
XX Human; target of methylation-induced silencing-1; TMS1; cytostatic;
KW antiproliferative; apoptosis inducer; gene therapy; CpG island;
KW caspase-recruiting domain; CARD; cancer; breast; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT misc_feature 1100..1725
FT /tag= a
FT /label= CpG island
FT /note= "This region is specifically claimed in claim 71"
FT CDS 1256..2361
FT /tag= b
FT /product= "Human TMS1 protein"
FT /note= "CDS is interrupted by 2 introns"
FT 1256..1529
FT exon

FT intron /tag= c
FT /number= 1
FT 1530..1742
FT /tag= d
FT /number= 1
FT 1743..1799
FT /tag= e
FT /number= 2
FT 1800..2104
FT /tag= f
FT /number= 2
FT 2105..2361
FT /tag= g
FT /number= 3
XX WO200129235-A2.
XX 26-APR-2001.
XX 18-OCT-2000; 2000WO-US28747.
XX 18-OCT-1999; 99US-0159975.
XX (UYEM-) UNIV EMORY.
XX Vertino PM;
XX WPI; 2001-290922/30.
XX P-PSDB; AAE00588.
XX Novel gene TMS1, transcriptionally silenced due to increased
XX methylation useful for identifying subject at risk of developing tumor
XX characterized by abnormal methylation, for treating cancer by inducing
XX apoptosis -
XX Claim 68; Page 112-113; 124pp; English.
XX The invention relates to identification of target of methylation-induced
XX silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to
XX abnormal methylation of a CpG island in its 5' regulatory region. TMS1
XX consists of a carboxy terminal caspase-recruiting domain (CARD) and
XX plays a role in induction of apoptosis. TMS1 gene and protein are useful
XX as tools for diagnosing and treating a subject at risk of developing
XX cancer (e.g. breast cancer) characterised by abnormal CpG methylation or
XX abnormally low levels of TMS1 expression products. Unique fragments of
XX TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.
XX TMS1 molecule is also useful for treating abnormal cell proliferation by
XX increasing TMS1 polypeptide level to an above normal level. The CpG
XX island region of TMS1 or its fragments are used to study the methylation
XX patterns apart from any coding region contained in it.
XX The present sequence is human target of methylation-induced silencing-1
XX (TMS1) genomic DNA.
SQ Sequence 2821 BP; 614 A; 864 C; 757 G; 586 T; 0 other;
Query Match 48.6%; Score 360; DB 22; Length 2821;
Best Local Similarity 100.0%; Pred. No. 5.6e-67;
Matches 360; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 381 CCAGGCGCTGCACCTTTATAGACCAGCACCGGGTGGCGTTATCCGAGGGTCAAAACGTT 440
Db 2101 CCAGGCGCTGCACCTTTATAGACCAGCACCGGGTGGCGTTATCCGAGGGTCAAAACGTT 2160
QY 441 GAGTGGCTGCTGGATGCTCTGTACGGAGGCTCTGTACGGATGAGCAGTACCGGAGTG 500
Db 2161 GAGTGGCTGCTGGATGCTCTGTACGGAGGCTCTGTACGGATGAGCAGTACCGGAGTG 2220
QY 501 CGGGCGGCGCCACCAACCAAGCAAGATGCGGAAGCTCTTCAGTTTTCACACGAGCTGG 560
Db 2221 CGGGCGGCGCCACCAACCAAGCAAGATGCGGAAGCTCTTCAGTTTTCACACGAGCTGG 2280
QY 561 AACTGAGCTTCAAGGAGCTTGTCTCCAGGCGCTTAAGGAGTCCAGGCTCTACCTGTTG 620
FT

Db	2281	AAC	TGGACCTGCAAGGACTTGCTCCTCCAGGCCCTAAGGGAGTCCCAAGTCCTACCTGGTG	2340
Qy	621	GAG	GACCTGGAGCGGAGCTGAGGCTCCTTTCCAGCAACACTCCGGTCAGCCCTGGCAAT	680
Db	2341	GAG	GACCTGGAGCGGAGCTGAGGCTCCTTTCCAGCAACACTCCGGTCAGCCCTGGCAAT	2400
Qy	681	CCC	ACCAATCATCCTGAAATCTGATCTTTTATACAAATATACGAAAAGCCAGCTTGAA	740
Db	2401	CCC	ACCAATCATCCTGAAATCTGATCTTTTATACAAATATACGAAAAGCCAGCTTGAA	2460

Search completed: January 29, 2004, 03:29:23
Job time : 184.487 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2004, 03:01:52 ; Search time 1474.73 Seconds
(without alignments)
12195.625 Million cell updates/sec

Title: US-09-996-617-7
Perfect score: 740
Sequence: 1 cgcgtccgctgcagcgggg.....atcagaaagccagcttgaa 740

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_hic:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_hic:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pln:*
- 20: em_gss_vit:*
- 21: em_gss_fun:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_rod:*
- 26: em_gss_pbg:*
- 27: em_gss_vrl:*
- 28: gb_gss1:*
- 29: gb_gss2:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	729.4	98.6	930	12	BM459678 AGENCOURT
2	727	98.2	804	12	BM926438 AGENCOURT
3	722	97.6	767	12	BM910918 AGENCOURT
4	721.2	97.5	792	12	BM549665 AGENCOURT

5	713.4	96.4	750	12	BG764161
c	704.8	95.2	725	14	CA309883
6	690.8	93.4	856	12	BG767422
7	688.6	93.1	748	10	BG337806
c	670.4	90.6	678	9	A1148558
10	667.6	90.2	741	12	BM051141
c	660	89.2	676	12	BM972785
11	654.2	88.4	675	10	BE908204
12	647.4	87.5	744	10	BG255521
13	647.4	87.5	744	10	BE909218
14	639	86.4	687	10	BE909218
c	617.2	83.4	638	12	BM998354
15	610	82.4	1079	12	BM456838
c	606	81.9	633	12	BM681310
17	604.6	81.7	1201	13	EX356773
18	581.2	78.5	830	10	BF238010
19	581.2	78.5	830	10	BF238010
c	553.8	74.8	621	12	BI468249
20	546	73.8	866	10	EG684386
c	544.6	73.6	634	9	A1991236
22	539	72.8	878	13	BU501178
23	536	72.4	888	10	BE906296
24	533	72.0	615	9	A1346818
c	511.4	69.1	542	12	BQ045282
26	488.4	66.0	518	13	RU626120
c	477.4	64.5	567	9	AA528254
28	471	63.6	572	14	CA439478
c	469.2	63.4	520	12	BI468250
30	468.4	63.3	486	12	BQ012745
c	463	62.6	500	10	AW973297
32	450.4	60.9	481	9	AA573948
c	446	60.3	446	12	BM726665
34	441.8	59.7	465	9	AA582937
c	431.6	58.3	700	14	CB429243
36	430.4	58.2	698	14	CB452677
37	426	57.6	695	14	CB422326
38	420.6	56.8	1093	12	BM920038
c	409.8	55.4	421	9	AA278825
40	408	55.1	671	14	CB421971
c	407	55.0	669	14	CB429601
42	390	52.7	565	12	BM362328
c	385.4	52.1	583	14	CB436146
44	382	51.6	395	9	A1262374
c					

ALIGNMENTS

RESULT 1
BM459678
LOCUS
DEFINITION
5', mRNA sequence.
AGENCOURT 6417789 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5534308
linear EST 05-FEB-2002

ACCESSION
BM459678
VERSION
BM459678.1
KEYWORDS
EST.
SOURCE
Homo sapiens (human)

ORGANISM
Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini, Hominidae; Homo.
1 (bases 1 to 930)
NIH-MGC http://mgc.ncl.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12220 row: j column: 05
High quality sequence stop: 670.

FEATURES	Location/Qualifiers		BM926438	
	source	1. .930	LOCUS	804 bp mRNA linear EST 12-MAR-2002
FEATURES	Location/Qualifiers		AGENCOURT 6644759 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5766806	
	source	1. .930	DEFINITION	5', mRNA sequence.
FEATURES	Location/Qualifiers		BM926438	
	source	1. .930	ACCESSION	BM926438
FEATURES	Location/Qualifiers		BM926438	
	source	1. .930	VERSION	1
FEATURES	Location/Qualifiers		BM926438	
	source	1. .930	KEYWORDS	EST.
FEATURES	Location/Qualifiers		BM926438	
	source	1. .930	SOURCE	Homo sapiens (human)
FEATURES	Location/Qualifiers		BM926438	
	source	1. .930	ORGANISM	Homo sapiens
FEATURES	Location/Qualifiers		BM926438	
	source	1. .930	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
FEATURES	Location/Qualifiers		BM926438	
	source	1. .930	AUTHORS	NIH-MGC http://mgc.nhl.nih.gov/
FEATURES	Location/Qualifiers		BM926438	
	source	1. .930	JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
FEATURES	Location/Qualifiers		BM926438	
	source	1. .930	COMMENT	Unpublished Contact: Robert Strausberg, Ph.D. Email: cga@remail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM12824 row: a column: 15 High quality sequence stop: 746.
FEATURES	Location/Qualifiers		BM926438	
	source	1. .930	FEATURES	Location/Qualifiers 1. .804 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:5766806" /lab_host="DH10B" /clone_lib="NIH_MGC_122" /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH MGC Library."
FEATURES	Location/Qualifiers		BM926438	
	source	1. .930	BASE COUNT	165 a 238 c 243 g 125 t 33 others
FEATURES	Location/Qualifiers		BM926438	
	source	1. .930	ORIGIN	Query Match Best Local Similarity 98.2%; Score 727; DB 12; Length 804; Matches 730; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
FEATURES	Location/Qualifiers		BM926438	
	source	1. .930	QY	5 TCCGGCTCAGCGGGTGGAGCGGGGAGCGCGGGGATCTGGAGCCATGGGCGCGC 64
FEATURES	Location/Qualifiers		BM926438	
	source	1. .930	Db	7 TCCCGGATAGCGGGGTGAGCGGGGAGCGCGGGGATCTGGAGCCATGGGCGCGC 66
FEATURES	Location/Qualifiers		BM926438	
	source	1. .930	QY	65 GCGGACCCATCTCTGGATGCTGGAGAACTGACCGCCGAGGAGCTCAAGAGTTCAA 124
FEATURES	Location/Qualifiers		BM926438	
	source	1. .930	Db	67 GCGGACCCATCTCTGGATGCTGGAGAACTGACCGCCGAGGAGCTCAAGAGTTCAA 126
FEATURES	Location/Qualifiers		BM926438	
	source	1. .930	QY	125 GCTGAAGTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 184
FEATURES	Location/Qualifiers		BM926438	
	source	1. .930	Db	127 GCTGAAGTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 186
FEATURES	Location/Qualifiers		BM926438	
	source	1. .930	QY	185 GCTGTCCATGGAGCGCTTGGACCTCAACGCAAGCTGTGTGTGTGTGTGTGTGTGT 244
FEATURES	Location/Qualifiers		BM926438	
	source	1. .930	Db	187 GCTGTCCATGGAGCGCTTGGACCTCAACGCAAGCTGTGTGTGTGTGTGTGTGTGT 246
FEATURES	Location/Qualifiers		BM926438	
	source	1. .930	QY	245 CGGCGCCGAGCTCAACGCTTAACGCTTAACGCTTAACGCTTAACGCTTAACGCTTA 304
FEATURES	Location/Qualifiers		BM926438	
	source	1. .930	Db	247 CGGCGCCGAGCTCAACGCTTAACGCTTAACGCTTAACGCTTAACGCTTAACGCTTA 306
FEATURES	Location/Qualifiers		BM926438	
	source	1. .930	QY	305 GCTGAGCGCGCCACGCAACGAGGCTCTGTGAGCGCGCCAGCTGGGATCCAGGCCCTCC 364

```

Db 307 GCTGCAGCGGCCACCCAGCGCTCTGGAGCCGCGCCAGCTGGGATCCAGGCCCTCC 366
QY 365 TCAGTCGGCAGCAGCAAGCCAGGCTGACACTTTATAGACCAGCACCAGGCTGCGCTTATCGC 424
Db 367 TCAGTCGGCAGCAGCAAGCCAGGCTGACACTTTATAGACCAGCACCAGGCTGCGCTTATCGC 426
QY 425 GAGGCTCACAACGTTGAGTGGCTGTGATGCTGTGTACGGGAAGGTCCTGACGGATGA 484
Db 427 GAGGCTCACAACGTTGAGTGGCTGTGATGCTGTGTACGGGAAGGTCCTGACGGATGA 486
QY 485 GCAGTACCAAGCAGTGGCGGCGAGCCACCAACCAAGCAAGATGCGGAAGCTTTCAG 544
Db 487 GCAGTACCAAGCAGTGGCGGCGAGCCACCAACCAAGCAAGATGCGGAAGCTTTCAG 546
QY 545 TTTCACACAGCCTGGAATCGGACCTGCAAGACTTGTCTCTCCAGGCCCTTAAGGAGTC 604
Db 547 TTTCACACAGCCTGGAATCGGACCTGCAAGACTTGTCTCTCCAGGCCCTTAAGGAGTC 606
QY 605 CCAGTCCTACTGTTGGAGGACCTGGAGCGGAGCTGAGGCTCTTCCAGCAACTCCG 664
Db 607 CCAGTCCTACTGTTGGAGGACCTGGAGCGGAGCTGAGGCTCTTCCAGCAACTCCG 666
QY 665 GTTCAGCCCTGCAATCCCAACCAATCATCTGAATCTGATCTTTTATACAAATATAC 724
Db 667 GTTCAGCCCTGCAATCCCAACCAATCATCTGAATCTGATCTTTTATACAAATATAC 726
QY 725 GAAAGCCAGCTTGAA 740
Db 727 GAAAGCCAGCTTGAA 742

RESULT 3
BM910918
LOCUS BM910918.1 GI:19361297
DEFINITION AGENCOURT_6615922 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5454497
5', mRNA sequence.
ACCESSION BM910918
VERSION BM910918.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L10M1949 row: 1 column: 18
High quality sequence stop: 641.
Location/Qualifiers
1..767
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5454497"
/tissue_type="astrocytoma grade IV, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_98"
/note="Organ: brain; Vector: pOTB7; Site: 1. XhoI; Site: 2.
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit"

```

FEATURES

source

```

RESULT 4
BM549665
LOCUS BM549665
DEFINITION AGENCOURT_6544142 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5745741
5', mRNA sequence.
ACCESSION BM549665
VERSION BM549665.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 163 a 241 c 237 g 124 t 2 others
ORIGIN
Query Match 97.6%; Score 722; DB 12; Length 767;
Best Local Similarity 99.7%; Pred. No. 3.8e-152;
Matches 733; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 7 CGGCTGCAGCGGGGTGAGCGCGGCGGCGGATCTTGGAGCCATGGGCGCGCGC 66
Db 17 CGGCTGCAGCGGGGTGAGCGCGGCGGCGGATCTTGGAGCCATGGGCGCGCGC 76
QY 67 GCGACGCCATCTTGGATGCGTGGAGAACTTACCGCGGAGAGTCAAGAGTTCAAGC 126
Db 77 GCGACGCCATCTTGGATGCGTGGAGAACTTACCGCGGAGAGTCAAGAGTTCAAGC 136
QY 127 TGAAGCTCTGCTGCGTGGCTGCGGAGGGGTACGGGCGCATCCGCGGGCGCGCTGC 186
Db 137 TGAAGCTCTGCTGCGTGGCTGCGGAGGGGTACGGGCGCATCCCGGGCGCGCTGC 196
QY 187 TGTCCATGAGCGCTTGGACCTACCGACAAGCTGCTCAGCTTCTACCTGGAGACTACG 246
Db 197 TGTCCATGAGCGCTTGGACCTACCGACAAGCTGCTCAGCTTCTACCTGGAGACTACG 256
QY 247 GCGCCGAGCTCACCGCTTAACGCTGCTGCGCGACATGGGCTTGCAGGAGATGGCGGCGCAGC 306
Db 257 GCGCCGAGCTCACCGCTTAACGCTGCTGCGCGACATGGGCTTGCAGGAGATGGCGGCGCAGC 316
QY 307 TGCAGGCGGCGACGACCAAGGCTTGGAGCGCGCGGCTTGGAGTCCAGGCGCTTCTCTC 366
Db 317 TGCAGGCGGCGACGACCAAGGCTTGGAGCGCGCGGCTTGGAGTCCAGGCGCTTCTCTC 376
QY 367 AGTCGCGCAGCAAGCCAGGCTTATAGACCAGCAGCGGCTGCGCTTATCGCGA 426
Db 377 AGTCGCGCAGCAAGCCAGGCTTATAGACCAGCAGCGGCTGCGCTTATCGCGA 436
QY 427 GGGTCACAAAAGCTTGGAGTGGCTGCTGTACGGGAAGGTCTCTGACGAGTACG 486
Db 437 GGGTCACAAAAGCTTGGAGTGGCTGCTGTACGGGAAGGTCTCTGACGAGTACG 496
QY 487 AGTACAGCGAGTGGCGGCGGAGCCACCAAGCAAGATGGCGGAGTCTTCAGTT 546
Db 497 AGTACAGCGAGTGGCGGCGGAGCCACCAAGCAAGATGGCGGAGTCTTCAGTT 556
QY 547 TCACACCAAGCTTGAAGTGGAGCTTCTCTCCAGGCTTCCAGGCGCTTAAAGGAGTCC 606
Db 557 TCACACCAAGCTTGAAGTGGAGCTTCTCTCCAGGCTTCCAGGCGCTTAAAGGAGTCC 616
QY 607 AGTCCTACTGTTGGAGGACCTTGGAGCGGAGTGGAGGCTTCTTCCAGCAACACTCCCGT 666
Db 617 AGTCCTACTGTTGGAGGACCTTGGAGCGGAGTGGAGGCTTCTTCCAGCAACACTCCCGT 676
QY 667 CAGCCCTTGGCAATCCCAAGAAATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 725
Db 677 CAGCCCTTGGCAATCCCAAGAAATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 736
QY 726 AAAAGCCAGCTTGAA 740
Db 737 AAAAGCCAGCTTGAA 751

```

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 792)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM12769 row: c column: 22
 High quality sequence stop: 570.
 Location/Qualifiers
 1..792
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5745741"
 /tissue_type="leukocyte"
 /lab_host="DH10B"
 /clone_lib="NIH MGC 118"
 /note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV
 (destroyed); RNA source leukocytes from anonymous pool of
 non-activated adult donors. Library is oligo-dT primed
 and directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb, insert size range
 1.2-3.3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 027. Note:
 this is a NIH MGC Library."
 182 a 241 c 242 g 122 t 5 others
 BASE COUNT 182 a 241 c 242 g 122 t 5 others
 ORIGIN
 Query Match 97.5%; Score 721.2; DB 12; Length 792;
 Best Local Similarity 98.8%; Pred. No. 5.7e-152;
 Matches 726; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 6 CCGCTGACGCGGGTGGAGCGGCGGCGGCGGATCCTGGAGCCATGGGCGCGG 65
 19 CTGGATACGCGGGTGGAGCGGCGGCGGCGGATCCTGGAGCCATGGGCGCGG 78
 66 CGGACGCCATCTCGATCGCTGGAGAACCTGACCGCGGAGAGCTCAAGATTCAAG 125
 79 CGGACGCCATCTCGATCGCTGGAGAACCTGACCGCGGAGAGCTCAAGATTCAAG 138
 126 CTGAAGCTGCTGCGGTGCGGTGCGGAGGCTACGGCGCATCCCGGCGGCGCTG 185
 139 CTGAAGCTGCTGCGGTGCGGTGCGGAGGCTACGGCGCATCCCGGCGGCGCTG 198
 186 CTGTCCATGGACGCTTGGACCTCACCAGACAGCTGGTCTACCTGGAGACCTAC 245
 199 CTGTCCATGGACGCTTGGACCTCACCAGACAGCTGGTCTACCTGGAGACCTAC 258
 246 GCGCGCGAGCTCACCCCTAACGTGCTGCGGACATGGGCTGCGAGAGATGCCGCGG 305
 259 GCGCGCGAGCTCACCCCTAACGTGCTGCGGACATGGGCTGCGAGAGATGCCGCGG 318
 306 CTGACGCGGCGGACCGACAGGGCTCTGGAGCGGCGGAGCTGGATCCAGGCCCTCT 365
 319 CTGACGCGGCGGACCGACAGGGCTCTGGAGCGGCGGAGCTGGATCCAGGCCCTCT 378
 366 CAGTCCGACGCAAGCAGCGCTTATAGACAGCAGCGGCTCGGCTTATCGCG 425
 379 CAGTCCGACGCAAGCAGCGCTTATAGACAGCAGCGGCTCGGCTTATCGCG 438
 426 AGGTCACAAAAGTTGAGTGGTCTGGATGCTCTGACGGAAGTCTGACGAGATGAG 485
 439 AGGTCACAAAAGTTGAGTGGTCTGGATGCTCTGACGGAAGTCTGACGAGATGAG 498

486 CAGTACCAGGAGTGGGCGGAGCCCAACCAAGCAAGATCGGAAGCTCTTCACT 545
 499 CAGTACCAGGAGTGGGCGGAGCCCAACCAAGCAAGATCGGAAGCTCTTCACT 558
 546 TTCACACAGGCTGGAACCTGACCTGGAAGACTTGTCTCCAGGCGCTTAAGGAGTCC 605
 559 TTCACACAGGCTGGAACCTGACCTGGAAGACTTGTCTCCAGGCGCTTAAGGAGTCC 618
 606 CAGTCTACCTGCTGGAGGAGCTGGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGG 665
 619 CAGTCTACCTGCTGGAGGAGCTGGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGG 678
 666 TCAGCCCTGCAATCCCAACCAATCATCTGAATCTGATCTTTTATACCAATATACG 725
 679 TCAGCCCTGCAATCCCAACCAATCATCTGAATCTGATCTTTTATACCAATATACG 738
 726 AAAAGCCAGCTTGAA 740
 739 AAAAGCCGAANAA 753
 RESULT 5
 BG764161 750 bp mRNA linear EST 15-MAY-2001
 LOCUS 602737190F1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4862182 5',
 DEFINITION mRNA sequence.
 ACCESSION BG764161
 VERSION BG764161.1 GI:14074814
 KEYWORDS EST.
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: ATCC/DCTD/DTF
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCMI722 row: d column: 23
 High quality sequence stop: 721.
 Location/Qualifiers
 1..750
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4862182"
 /tissue_type="melanotic melanoma, high MDR (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC_49"
 /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; CDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 Library."|
 BASE COUNT 165 a 232 c 235 g 118 t
 ORIGIN
 Query Match 96.4%; Score 713.4; DB 12; Length 750;
 Best Local Similarity 99.7%; Pred. No. 3.2e-150;
 Matches 725; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

15 GCGGGGTGAGCGCGGAGCGCGCGGGGATCTGAGGACCATGGGCGCGCGCGGACGCC 74
 Db GCGGGGTGAGCGCGGAGCGCGCGGGGATCTGAGGACCATGGGCGCGCGGACGCC 61
 75 ATCCTGGATGCGCTGGAGAACCTGACCGCGGAGGAGCTCAAGACTTCAAGCTGAAGCTG 134
 Db ATCCTGGATGCGCTGGAGAACCTGACCGCGGAGGAGCTCAAGACTTCAAGCTGAAGCTG 121
 135 CTGTGCGTGGCGCTGCGGAGGCGCTACGGGCGCATCCGCGGGGCGCGCTGTGCTCCATG 194
 Db CTGTGCGTGGCGCTGCGGAGGCGCTACGGGCGCATCCGCGGGGCGCGCTGTGCTCCATG 181
 195 GACGCTTGGACCTCACCGACAAAGCTGTGCTTACCTGGAGACCTACGGCGCGGAG 254
 Db GACGCTTGGACCTCACCGACAAAGCTGTGCTTACCTGGAGACCTACGGCGCGGAG 241
 255 CTCACCGCTAAGCTGTGCGGACATGGGCTGCGAGGATGGCGGCGAGCTGCGAGGG 314
 Db CTCACCGCTAAGCTGTGCGGACATGGGCTGCGAGGATGGCGGCGAGCTGCGAGGG 301
 315 GCCAGCGCACCGGGCTCTGGAGCGCGCGCAGCTGGGATCCAGGCGCCCTCTCAGTCGGCA 374
 Db GCCAGCGCACCGGGCTCTGGAGCGCGCGCAGCTGGGATCCAGGCGCCCTCTCAGTCGGCA 361
 375 GCCAAGCCAGGCTGACCTTTATAGACAGCACCGGCTGCGCTTATCGGAGGTCACA 434
 Db GCCAAGCCAGGCTGACCTTTATAGACAGCACCGGCTGCGCTTATCGGAGGTCACA 421
 435 AACGTTGAGTGTGCTGTGATCTCTGTACGGGAGGTCCTACGGATGAGCAGTACCAG 494
 Db AACGTTGAGTGTGCTGTGATCTCTGTACGGGAGGTCCTACGGATGAGCAGTACCAG 481
 495 GCAGTGGCGGCGGAGCGCACCAACCAAGAGTGGGAGCTCTTCAAGTTTACACCA 554
 Db GCAGTGGCGGCGGAGCGCACCAACCAAGAGTGGGAGCTCTTCAAGTTTACACCA 541
 555 GCGTGAACCTGACCTGCAAGGACTTGTCTCCAGCGCCCTAAGGAGTCCCGAGTCTAC 614
 Db GCGTGAACCTGACCTGCAAGGACTTGTCTCCAGCGCCCTAAGGAGTCCCGAGTCTAC 601
 615 CTGGTGGAGGACTGGAGCGGAGCTGAGGCTCTTCCGAGCAACACTCCGGTCAAGCCCT 674
 Db CTGGTGGAGGACTGGAGCGGAGCTGAGGCTCTTCCGAGCAACACTCCGGTCAAGCCCT 661
 675 GCAATCCCAACCAATCATCTGATCTCTTTTATACA-CAATATACGAAGAAGCA 733
 Db GCAATCCCAACCAATCATCTGATCTCTTTTATACAGCAATATACGAAGAAGCA 721
 734 GCTTGAA 740
 Db GCTTGGA 728

RESULT 6
 CA309883/c
 LOCUS
 DEFINITION
 UI-H-F11-bid-e-07-0-UI.s1 NCI CGAP F11 Homo sapiens cDNA clone
 UI-H-F11-bid-e-07-0-UI 3', mRNA sequence.
 CA309883
 VERSION
 CA309883.1 GI:24472937
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 725)
 REFERENCE
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 AUTHORS
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Dr. Gary W. Hunninghake, U of I

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained
 from Dr. M. Bento Soares, bento-soares@uiowa.edu
 Seq primer: M13 FORWARD
 POLYA=yes.

FEATURES source

Location/Qualifiers
 1..725
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clones="UI-H-F11-bid-e-07-0-UI"
 /tissue_type="Aveolar Macrophage"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP F11"
 /note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
 modified polylinker; Site1: EcoR I; Site2: Not I;
 NCI CGAP F11 is a normalized cDNA library constructed from
 a pool of 81 RNA samples from Alveolar Macrophages
 challenged with different treatments. The library was
 normalized according to Bonaldo, Lennon and Soares, Genome
 Research, 6:791-806, 1996. First strand cDNA synthesis was
 primed with an oligo-dT primer containing a Not I site.
 Double stranded cDNA was ligated to an EcoR I adaptor,
 digested with Not I, and cloned directionally into
 pT73-Pac vector. The oligonucleotide used to prime the
 synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (dT)18 tail. The sequence tag for this library is
 GGCATATGCCG. The tissue was provided by Dr. Gary W.
 Hunninghake of the University of Iowa.
 TAG LIB-UI-H-F11
 TAG-TISSUE=Human Lung Aveolar Macrophage
 TAG_SEQ=GGCCATGCCG"

BASE COUNT 119 a 221 c 226 g 159 t
 ORIGIN
 Query Match 95.2%; Score 704.8; DB 14; Length 725;
 Best Local Similarity 99.7%; Pred. No. 2.7e-148;
 Matches 706; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 33 GCGGCGGGGATCTCTGGAGCCTACGGGCGCGCGCGGCGGAGCTCCTGTGCTGGAG 92
 Db 724 GCGGCGGGGATCTCTGGAGCCTACGGGCGCGCGCGGCGGAGCTCCTGTGCTGGAG 665
 QY 93 AACCTGACCGCGGAGGAGCTCAAGAAAGTTCAAGCTGAAGCTGCTGTGCTGCGCTGGC 152
 Db 664 AACCTGACCGCGGAGGAGCTCAAGAAAGTTCAAGCTGAAGCTGCTGTGCTGCGCTGGC 605
 QY 153 GAGGGCTACGGGCGCATCCCGGGGCGGCTGTCTCCATGGAGCGCTTGGACCTCAC 212
 Db 604 GAGGGCTACGGGCGCATCCCGGGGCGGCTGTCTCCATGGAGCGCTTGGACCTCAC 545
 QY 213 GACAGCTGTGTCAGCTTCTTACCTGGAGACCTACCGCGCGGAGCTCACCGTAACTGCTG 272
 Db 544 GACAGCTGTGTCAGCTTCTTACCTGGAGACCTACCGCGCGGAGCTCACCGTAACTGCTG 485
 QY 273 GCGCATATGGGCTTCAGGAGATGGCGGCGGAGCTGTCAGGGGCGGCAACGACACAGGGCTCT 332
 Db 484 GCGCATATGGGCTTCAGGAGATGGCGGCGGAGCTGTCAGGGGCGGCAACGACACAGGGCTCT 425
 QY 333 GGAGCGGCGGAGCTGGGATCCAGGCGGCTCTCTAGTCGGGAGCCAGCCAGGCGCTGCAC 392
 Db 424 GGAGCGGCGGAGCTGGGATCCAGGCGGCTCTCTAGTCGGGAGCCAGCCAGGCGCTGCAC 365
 QY 393 TTTATAGACCAACACCGGGCTGCGCTTATCGGAGGGGTCAACAAAGTTGAGTGGCTGCTG 452
 Db 364 TTTATAGACCAACACCGGGCTGCGCTTATCGGAGGGGTCAACAAAGTTGAGTGGCTGCTG 305
 QY 453 GATGCTCTGTACGGGAAGGTCTCTGACGATGAGCAGTACACGAGGAGTGGCGGCCAGGCC 512


```
FEATURES
source
High quality sequence stop: 710.
Location/Qualifiers
1..748
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4553378"
/tissue_type="leiomyosarcoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 46"
/note="Organ: uterus; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
BASE COUNT 153 a 239 c 238 g 118 t
ORIGIN
Query Match 93.1%; Score 688.6; DB 10; Length 748;
Best Local Similarity 98.5%; Pred. No. 1.2e-144; Indels 2; Gaps 2;
Matches 716; Conservative 0; Mismatches 9;
QY 15 GCGGGGTGAGCGCGGACGCGCGGGGATCTGGAGCCATGGGCGCGCGACGCG 74
DB 2 GCGGGGTGAGCGCGGACGCGCGGGGATCTGGAGCCATGGGCGCGCGACGCG 61
QY 75 ATCTGGATGCGCTGGAGAACTGACCGCGGAGGAGCTCAAGAAGTTCAAGCTGAAGCTG 134
DB 62 ATCTGGATGCGCTGGAGAACTGACCGCGGAGGAGCTCAAGAAGTTCAAGCTGAAGCTG 121
QY 135 CTGTGGTGGCTGCGGAGGCTGACGGCGCATCCCGGGGCGCGTGTCTTCATG 194
DB 122 CTGTGGTGGCTGCGGAGGCTGACGGCGCATCCCGGGGCGCGTGTCTTCATG 181
QY 195 GACGCTTGGACCTCACCACAACTGGTTCAGCTTCTACTGAGACCTACCGCGCGGAG 254
DB 182 GACGCTTGGACCTCACCACAACTGGTTCAGCTTCTACTGAGACCTACCGCGCGGAG 241
QY 255 CTCACGCTTAACTGCTGCGGACATGGGCTTCAGAGAGATGGCGGGGAGTGCAGGGG 314
DB 242 CTCACGCTTAACTGCTGCGGACATGGGCTTCAGAGAGATGGCGGGGAGTGCAGGGG 301
QY 315 GCCAGCACAGGGCTCTGGAGCGCGGACGCTGGGATCAGGCGGCTCTCAGTCGGCA 374
DB 302 GACAGCACAGGGCTCTGGAGCGCGGACGCTGGGATCAGGCGGCTCTCAGTCGGCA 361
QY 375 GCCAGCACAGGGCTCTGGAGCGCGGACGCTGGGATCAGGCGGCTCTCAGTCGGCA 434
DB 362 GCCAGCACAGGGCTCTGGAGCGCGGACGCTGGGATCAGGCGGCTCTCAGTCGGCA 421
QY 435 AACGTTGAGTGGCTGCTGGATGCTCTGTACGGAAGGTCCTGACGAGTACGAGTACAG 494
DB 422 AACGTTGAGTGGCTGCTGGATGCTCTGTACGGAAGGTCCTGACGAGTACGAGTACAG 481
QY 495 GCAGTGGGGGCGAGCCCAACACCAAGAGATCGGAAGCTCTTCAAGTTTACACCA 554
DB 482 GCAGTGGGGGCGAGCCCAACACCAAGAGATCGGAAGCTCTTCAAGTTTACACCA 541
QY 555 GCTGGAACTGGACCTGCAAGGACTTGTCTTCAGGCGCTAAGGAGTCCAGTCTCTAC 614
DB 542 GCTGGAACTGGACCTGCAAGGACTTGTCTTCAGGCGCTAAGGAGTCCAGTCTCTAC 601
QY 615 CTGGTGGAGGACCTGGAGCGGAGTCTGAGGCTCTCTCCAGCAACACTCCGGTTCAGCC 673
DB 602 CTGGTGGAGGACCTGGAGCGGAGTCTGAGGCTCTCTCCAGCAACACTCCGGTTCAGCC 661
QY 674 TGGCAATCCACCAAAATCATCTTGAATCTGATCTTTTATATACAAATATACGAAGGCA 733
DB 662 TGTCAATCCAGCAGATCATCTTGAATCTGATCTTTTATATACAAATATACGAAGGCA 720
```

```
QY 734 GCTTGAA 740
DB 721 GCTTGCA 727

RESULT 9
AII48558/c
LOCUS
DEFINITION
AII48558 678 bp mRNA linear EST 28-OCT-1998
qc68g10.xl Soares placenta 8to9weeks 2NBHP8to9W Homo sapiens cDNA
clone IMAGE:1714818 3' similar to TR015553 O15553 PYRIN. ;, mRNA
sequence.
AII48558
AII48558.1 GI:3677027
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo...
REFERENCE
1 (bases 1 to 678)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 804 Std Error: 0.00
Seq Primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 402.
Location/Qualifiers
1..678
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1714818"
/dev_stage="two placentae: one from 8 weeks and another
from 9 weeks post conception"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares_placenta_8to9weeks_2NBHP8to9W"
/note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTACCAATCTGAAGTGGAGCGCGCGGATTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library constructed by Bento Soares and
M.Fatima Bonaldo."
BASE COUNT 114 a 202 c 217 g 145 t
ORIGIN
Query Match 90.6%; Score 670.4; DB 9; Length 678;
Best Local Similarity 99.9%; Pred. No. 1.5e-140;
Matches 671; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 69 GACGCCATCTCGATGCGCTGGAGAACTGCGCGGAGGAGCTCAAGAAGTTCAAGCTG 128
DB 678 GACGCCATCTCGATGCGCTGGAGAACTGCGCGGAGGAGCTCAAGAAGTTCAAGCTG 619
QY 129 AAGCTGTGTGCTGGTGGCTGGCGCGAGGCGCATCCCGCGGCGCGCTGTGCTG 188
DB 618 AAGCTGTGTGCTGGTGGCTGGCGCGAGGCGCTACGGGCGCATCCCGCGGCGCGCTGCTG 559
QY 189 TCATGGAGCGCTTGGAGCTCAACGACAGTGTGTGCTTCTACTGAGACCTACGCG 248
DB 558 TCATGGAGCGCTTGGAGCTCAACGACAGTGTGTGCTTCTACTGAGACCTACGCG 499
QY 249 GCCGAGCTCACCGCTAAAGTGTGTGGCAGCATGGGCGCTGCAGAGATGCCCGGCGAGCTG 308
DB 498 GCCGAGCTCACCGCTAAAGTGTGTGGCAGCATGGGCGCTGCAGAGATGCCCGGCGAGCTG 439
```


[illegible][illegible]

ACCESSION	BM051141
VERSION	BM051141.1
KEYWORDS	GI:16780408
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE	1 (bases 1 to 741)
JOURNAL	NIH-MGC http://mgc.ncbi.nlm.nih.gov/ .
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished Contact: Robert Strausberg, Ph.D.

FEATURES source

RESULT 11	ACCESSION
BM972785/c	VERSION
LOCUS	KEYWORDS
DEFINITION	SOURCE

in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library. | "

Query Match	90.2%	Score 667.6	DB 12	Length 741
Best Local Similarity	97.1%	Pred. No. 6.4e-140		
Matches 701	Conservative 0	Mismatches 19	Indels 2	Gaps 2
QY	7	CGGCTGCAGCGGGGTGAGCGGGCGAGCGCGCGGGGATCTGGAGCCATGGGGCGCGCG	66	
DB	21	CGGCTGCAGCGGGGTGAGCGGGCGAGCGCGCGGGATCTTGAGCCATGGGGCGCGCG	80	
QY	67	GCAGCGCATCTCGGATCGCTGGAGAACTGACCGCGGAGGAGCTCAAGAAAGTTCAAGC	126	
DB	81	GCAGCGCATCTCGGATCGCTGGAGAACTGACCGCGGAGTGTCAAGAAAGTTCAAGC	140	
QY	127	TGAAGCTGCTGTCGGGTGCGCTCGGAGAGGCTACCGGCGCATCCCGGGGCGCGGTGC	186	
DB	141	TGAAGCTGCTGTCGGGTGCGCTCGGAGAGGCTACCGGCGCATCCCGGGGCGCGGTGC	200	
QY	187	TGTTCATGGAGCGCTTGGAGCTCACCGACAAGGTGGTCAGCTTCTACTTGAGACCTACG	246	
DB	201	TGTTCATGGAGCGCTTGGAGCTCACCGACAAGGTGGTCAGCTTCTACTTGAGACCTACG	260	
QY	247	GCSCCGAGCTCAACGCTTAAGCTGTCGGACATATGGGCTTGAGGAGATGCGCGGGGAGC	306	
DB	261	GCSCCGAGCTCAACGCTTAAGCTGTCGGACATATGGGCTTGAGGAGATGCGCGGGGAGC	320	
QY	307	TGCAGCGGGCCAGCCACGAGGGCTCTGGAGCGCGCCAGCTGGGATCCAGSCCCTCTCTC	366	
DB	321	TGCAGCGGGCCAGCCACGAGGGCTCTGGAGCGCGCCAGCTGGGATCCAGSCCCTCTCTC	380	
QY	367	AGTCGGCAGCCAAAGCCAGGCTCGACCTTTATAGACACCGGGCTGCGCTTATCGCGA	426	
DB	381	AGTCGGCAGCCAAAGCCAGGCTCGACCTTTATAGACACCGGGCTGCGCTTATCGCGA	440	
QY	427	GGGTCAAAACGTTGAGTGGCTGCTGGATGCTCTGTACGGGAAGGTCTTGACGGATGAGC	486	
DB	441	GGGTCAAAACGTTGAGTGGCTGCTGGATGCTCTGTACGGGAAGGTCTTGACGGATGAGC	500	
QY	487	AGTACCAAGGAGTGCAGGCGGAGCCACCAACCCAAGCAAGATGCGGAAGCTCTTCAAGTT	546	
DB	501	AGTACCAAGGAGTGCAGGCGGAGCCACCAACCCAAGCAAGATGCGGAAGCTCTTCAAGTT	560	
QY	547	TCACACGAGCTTGGAACTGGACCTGCAAGGATTTGCTCTCCAGGCGCTTAAGGGAGTCCC	606	
DB	561	TCACACGAGCTTGGAACTGGACCTGCAAGGATTTGCTCTCCAGGCGCTTAAGGGAGTCCC	620	
QY	607	AGTCCTACCTGGTGGAGGACCT-GGAGCGAGAGTGGGCTCTTCCAGACGACACTCCGG	665	
DB	621	AGTCCTACCTGGTGGAGGACCTGGAGCGGGAATGAGGCTCTTCCAGACGACACTCCGG	680	
QY	666	TCAGGCCCTGGCAATCCACCAATCATCTCGAATCTGATCTTTTATACACAAATATACG	725	
DB	681	TCAGGCCCTGGCAAT-CCACCAATCATCTCGAATCTGATCTTCTCATACCCCACTCCG	739	
QY	726	AA 727		
DB	740	AA 741		

BM972785 676 bp mRNA linear EST 20-FEB-2003
 UI-CF-EC1-abr-k-24-0-UI.s1 UI-CF-EC1 Homo sapiens
 UI-CF-EC1-abr-k-24-0-UI 3', mRNA sequence.
 BM972785
 BM972785.1 GI:19590376
 EST.
 Homo sapiens (human)

4

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 676)
 AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: McCray, PB
 McCray Lab
 University of Iowa
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
 Tel: 319 356 4866
 Fax: 319 356 7171
 Email: paul-mccray@uiowa.edu
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa
 cDNA library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.reagen.com) or from Open Biosystems
 (www.openbiosystems.com).
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES
 Location/Qualifiers
 1..676
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-CF-EC1-abr-k-24-0-UI"
 /tissue_type="Lung"
 /dev_stage="Adult and Fetal"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-CF-EC1"
 /note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
 modified polylinker; Site 1: EcoR I; Site 2: Not I;
 UI-CF-EC1 is a normalized cDNA library containing the
 following tissue(s): Normal lung from adult and from fetal
 day 64, day 87, week 19 and week 42. The library was
 constructed according to Bonaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an EcoR I
 adaptor, digested with Not I, and cloned directionally
 into pT73-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (dT)18 tail. The sequence tag for this library is
 AAGTGTCTAC.
 TAG_LIB=UI-CF-EC1
 TAG_TISSUE=Normal Lung Epithelial Cells Tissue nos 369-371
 and 380-383
 TAG_SEQ=AAGTGTCTAC"

BASE COUNT 112 a 199 c 211 g 154 t
 ORIGIN

Query Match 89.2%; Score 660; DB 12; Length 676;
 Best Local Similarity 100.0%; Pred. No. 3.2e-138;
 Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 GATCGCTGGAGAACCTGACCGCGGAGGAGCTCAAGAGTTCAAGCTGAAGCTGTGTGCG 140
 DB 676 GATCGCTGGAGAACCTGACCGCGGAGGAGCTCAAGAGTTCAAGCTGAAGCTGTGTGCG 617
 QY 141 GTCCCGCTGCGGAGGCTACGGCGGCATCCCGGGGGCGCGTGTGTCTTCATGGACGCC 200
 DB 616 GTCCCGCTGCGGAGGCTACGGCGGCATCCCGGGGGCGCGTGTGTCTTCATGGACGCC 557
 QY 201 TTGGACCTCACCAGCAAGCTGTGTACCTTCTACCTGGAGACCTACCGGCCCGAGCTCAC 260
 DB 556 TTGGACCTCACCAGCAAGCTGTGTACCTTCTACCTGGAGACCTACCGGCCCGAGCTCAC 497

QY 261 GCTAACGTGTGCGCGAATGGGCTGCAGAGATGGCCGGGAGCTGCAGGGGCGCCACG 320
 DB 496 GCTAACGTGTGCGCGAATGGGCTGCAGAGATGGCCGGGAGCTGCAGGGGCGCCACG 437
 QY 321 CACCAGGGCTTGAGCGCGCCAGCTGGGATCCAGGCCCTCTCTCAGTCCGGCAGCCCAAG 380
 DB 436 CACCAGGGCTTGAGCGCGCCAGCTGGGATCCAGGCCCTCTCTCAGTCCGGCAGCCCAAG 377
 QY 381 CCAGGCTGCACCTTTATAGACACAGCACCGGCTGGCTTATCGCGAGGGTCACAAACTT 440
 DB 376 CCAGGCTGCACCTTTATAGACACAGCACCGGCTGGCTTATCGCGAGGGTCACAAACTT 317
 QY 441 GAGTGGCTGTGGATGCTCTGTACGGGAAGTCTTCACGGATGACAGTACAGGACGTG 500
 DB 316 GAGTGGCTGTGGATGCTCTGTACGGGAAGTCTTCACGGATGACAGTACAGGACGTG 257
 QY 501 CGGGCGAGGCCACCAACCCAGCAAGATCGGAGCTCTTCAGTTTACACACCGCCCTGG 560
 DB 256 CGGGCGAGGCCACCAACCCAGCAAGATCGGAGCTCTTCAGTTTACACACCGCCCTGG 197
 QY 561 AACTGACCTGCAAGGACTTGTCTCTCCAGGCCCTTAAGGGAGTCCCACTTCTACCTGGTG 620
 DB 136 AACTGACCTGCAAGGACTTGTCTCTCCAGGCCCTTAAGGGAGTCCCACTTCTACCTGGTG 137
 QY 621 GAGGACCTGGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGCTCAGCCCTGGCAAT 680
 DB 136 GAGGACCTGGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGCTCAGCCCTGGCAAT 77
 QY 681 CCACCAATCATCTCTGAATCTGATCTTTTATACAAATATACGAAAGCCAGCTTGA 740
 DB 76 CCACCAATCATCTCTGAATCTGATCTTTTATACAAATATACGAAAGCCAGCTTGA 17

RESULT 12
 BE908204 LOCUS
 DEFINITION BE908204 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3902319 5',
 mRNA sequence.
 ACCESSION BE908204
 VERSION BE908204.1 GI:10402543
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 675)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM9704 row: j column: 16
 High quality sequence stop: 660.

FEATURES
 Location/Qualifiers
 1..675
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3902319"
 /tissue_type="epithelioid carcinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_70"
 /note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.1 kb. Library constructed by Life

Technologies."

BASE COUNT 151 a 214 c 198 g 112 t

Query Match 88.4%; Score 654.2; DB 10; Length 675;
 Best Local Similarity 99.5%; Pred. No. 6.5e-137;
 Matches 656; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 82 ATGCGCTGAGAACCTGACCGCGAGGAGCTCAAGAAGTTCAAGCTGGAAGCTGCTGCGG 141
 DB 1 ATGCGCTGAGAACCTGACCGCGAGGAGCTCAAGAAGTTCAAGCTGGAAGCTGCTGCGG 60

QY 142 TGCCTGTCGCGCAGGAGCTACGCGCGCATCCCGCGGGCGCGTGTCTGTCATGAGACGCT 201
 DB 61 TGCCTGTCGCGCAGGAGCTACGCGCGCATCCCGCGGGCGCGTGTCTGTCATGAGACGCT 120

QY 202 TGGACCTACCGAACAGCTGTGTCAGCTTCTACTGTGAGACCTACGCGCGCGAGCTCACCG 261
 DB 121 TGGACCTACCGAACAGCTGTGTCAGCTTCTACTGTGAGACCTACGCGCGCGAGCTCACCG 180

QY 262 CTAAGCTGCTGCGGACATGGGCTGCGAGAGATGCGCGGCGAGCTGCGAGGCGGCGACGC 321
 DB 181 CTAAGCTGCTGCGGACATGGGCTGCGAGAGATGCGCGGCGAGCTGCGAGGCGGCGACGC 240

QY 322 ACCAGGGCTCTGGAGCGCGCCAGCTGGGATCCAGGCCCTCTCTCAGTCGGCAGCAAGC 381
 DB 241 ACCAGGGCTCTGGAGCGCGCCAGCTGGGATCCAGGCCCTCTCTCAGTCGGCAGCAAGC 300

QY 382 CAGGGCTGCATTTATAGACAGACCGCGGCTGCGCTTATCGCGAGGFTCAAAACGTTG 441
 DB 301 CAGGGCTGCATTTATAGACAGACCGCGGCTGCGCTTATCGCGAGGFTCAAAACGTTG 360

QY 442 AGTGGCTGCTGAGTCTCTGTACGGAAGTCTTACCGATGAGCAGTACACAGGCGAGTGC 501
 DB 361 AGTGGCTGCTGAGTCTCTGTACGGAAGTCTTACCGATGAGCAGTACACAGGCGAGTGC 420

QY 502 GGGCGGAGCCCAACCAAGCAAGATCGGGAAGCTCTTTCAGTTTCAACACAGCTTGA 561
 DB 421 GGGCGGAGCCCAACCAAGCAAGATCGGGAAGCTCTTTCAGTTTCAACACAGCTTGA 480

QY 562 ACTGAGCTGCAAGGACTTGTCTTCCAGGCGCTTAAGGAGTCCAGTCTTACCTGTTGG 621
 DB 481 ACTGAGCTGCAAGGACTTGTCTTCCAGGCGCTTAAGGAGTCCAGTCTTACCTGTTGG 540

QY 622 AGGACTGAGGCGGAGCTGAGCTCTCTCCAGCAACACTCCGTCAGCCCTGGCAATC 681
 DB 541 AGGACTGAGGCGGAGCTGAGCTCTCTCCAGCAACACTCCGTCAGCCCTGGCAATC 600

QY 682 CCACCAATCATCTGAACTGATCTTTTATACAAATATACGAAAGCCAGCTTGA 740
 DB 601 CCACCAATCATCTGAACTGATCTTTTATACAAATATACGAAAGCCAGCTTGA 659

RESULT 13
 BG255521
 LOCUS 744 bp mRNA linear EST 13-PEB-2001
 DEFINITION 602367671F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4476088 5', mRNA sequence.

ACCESSION BG255521
 VERSION BG255521.1 GI:12765259
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 744)
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: DCTD/DTF
 cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLAM10302 row: m column: 17
 High quality sequence stop: 661.
 Location/Qualifiers
 1..744
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clones="IMAGE:4476088"
 /tissue_type="adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_91"
 /note="Organ: prostate; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.4 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
 BASE COUNT 159 a 227 c 238 g 120 t

FEATURES
 source
 Query Match 87.5%; Score 647.4; DB 10; Length 744;
 Best Local Similarity 97.0%; Pred. No. 2.2e-135;
 Matches 713; Conservative 0; Mismatches 16; Indels 6; Gaps 5;

QY 8 GGTGTCAGCGGGTGTAGCGCGGCGAGCGCGCGGGATCTCTGAGCGCATGGCGCGCGCGCG 67
 DB 1 GGTGTCAGCGGGTGTAGCGCGGCGAGCGCGCGGGATCTCTGAGCGCATGGCGCGCGCGCG 60

QY 68 CGAGCCATCTCGATGCGCTGGAGAACCTCACCGCGGAGGAGCTCAAGAAAGTTCAAGCT 127
 DB 61 CGAGCCATCTCGATGCGCTGGAGAACCTCACCGCGGAGGAGCTCAAGAAAGTTCAAGCT 120

QY 128 GAAGTGTGTGTCGTCGCGGAGGAGTACGGGCGCATCCCGCGGGCGCGCTGCT 187
 DB 121 GAAGTGTGTGTCGTCGCGGAGGAGTACGGGCGCATCCCGCGGGCGCGCTGCT 180

QY 188 GTCCATGAGCGCTTGGACCTCACGCAAGCTGTGAGCTTCTACCTGGAGACCTACGG 247
 DB 181 GTCCATGAGCGCTTGGACCTCACGCAAGCTGTGAGCTTCTACCTGGAGACCTACGG 240

QY 248 CGCGAGCTCACCGCTAACGTCGCGGACATGGGCTGCGAGAGTGGCGGGGCGCGAGCT 307
 DB 241 CGCGAGCTCACCGCTAACGTCGCGGACATGGGCTGCGAGAGTGGCGGGGCGCGAGCT 300

QY 308 GCAGCGGCCACGCGACCGAGGCTCTGGAGCGCGCGAGCTGGGATCCAGGCCCTCTCTCA 367
 DB 301 GCAGCGGCCACGCGACCGAGGCTCTGGAGCGCGCGAGCTGGGATCCAGGCCCTCTCTCA 360

QY 368 GTCGGCAGCCAAAGCGAGCGCTTGTATATAGACAGCACCGGGCTGGCGTTATGCGGAG 427
 DB 361 GTCGGCAGCCAAAGCGAGCGCTTGTATATAGACAGCACCGGGCTGGCGTTATGCGGAG 420

QY 428 GGTACAAAACCTTCAGTGGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 487
 DB 421 GGTACAAAACCTTCAGTGGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480

QY 488 GTACCCAGGCTGCGGGCGGAGCGCCACCAACCAAGCAAGATGGCGAAGCTCTTCAGTTT 547
 DB 481 GTACCCAGGCTGCGGGCGGAGCGCCACCAACCAAGCAAGATGGCGAAGCTCTTCAGTTT 539

QY 548 CACACAGCTGGAACCTGGAACCTGGAACCTGGAACCTGGAACCTGGAACCTGGAACCTGGA 607
 DB 540 CACACAGCTTGAACCTGGAACCTGGAACCTGGAACCTGGAACCTGGAACCTGGAACCTGGA 597

QY 608 GTCCTTACCTGTCGAGGAGCTGGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGCTC 667
 DB 598 GTCCTTACCTGTCGAGGAGCTGGAGCGGAGCTGAGGCTCTTCCAGCAACACTCCGCTC 656

QY 668 AGCCCTTGGCAATCCCAACCAATCATCTCTGAAT--CTGATCTTTTATACAAATATACG 725

```

Db      657  AGCCCCGGGAAATCCACAAAATATCTCTGGATTCTGATCTTTTATACAAATATAGG 716
      726  AAAAGCCAGCTTGAA 740
      717  AAAAGCCAGTTGAAA 731

RESULT 14
BE909218
LOCUS   BE909218
DEFINITION BE909218
ACCESSION BE909218
VERSION   BE909218
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 687)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished
COMMENT    Contact: Robert Strausberg, Ph.D.
           Email: cgabbs-r@mail.nih.gov
           Tissue Procurement: ATCC
           CDNA Library Preparation: Life Technologies, Inc.
           CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM9707 row: k column: 19
           High quality sequence stop: 673.
           Location/Qualifiers
             1..687
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:3903498"
               /tissue_type="epithelioid carcinoma"
               /lab_host="DH10B (phage-resistant)"
               /clone_lib="NIH MGC 70"
               /note="Organ: pancreas; Vector: pCMV-SPORT6; Site 1: Not I;
               Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
               Average insert size 1.1 Kb. Library constructed by Life
               Technologies."

BASE COUNT 140 a 219 c 214 g 114 t
ORIGIN

Query Match 86.4%; Score 639; DB 10; Length 687;
Best Local Similarity 98.7%; Pred.No. 1.7e-133;
Matches 676; Conservative 0; Mismatches 5; Indels 4; Gaps 3;

QY 48 GGAGCCATGGGGCGCG-CGGCGAGCCCATCTGTGATGGCTGGAGAACTGACGCGCA 106
Db 1 GGAGCCATGGGGCGCGCGCGCGAGCCCATCTGTGATGGCTGGAGAACTGACGCGCA 60

QY 107 GGAGCTCAAGAAGTTCAAGCTGAAGCTGTGTGGTGGCGCTGGCGAGGGCTACGGGCG 166
Db 61 GGAGCTCAAGAAGTTCAAGCTGAAGCTGTGTGGTGGCGCTGGCGAGGGCTACGGGCG 120

QY 167 CATCCCGGGGGCGCGCTGTGTCATGACCGCTTGGACCTCACCAGCAAGCTGTGTAG 226
Db 121 CATCCCGGGGGCGCGCTGTGTCATGACCGCTTGGACCTCACCAGCAAGCTGTGTAG 180

QY 227 TTCTACTGGAGACCTACGGCGCGAGCTCACCGCTAAGCTGCTCGCGAGCATGGGCT 286
Db 181 TTCTACTGGAGACCTACGGCGCGAGCTCACCGCTAAGCTGCTCGCGAGCATGGGCT 240

QY 287 GCAGGAGATGGCGGGCGAGCTGCAGGGCGGCGACCGACCGAGGGCTGTGGAGCGCGCCAGC 346

```

```

Db      241  GCAGGAGATGGCGGGGAGCTGCGAGGCGGCCACGACAGGGCTCTGGAGCGGCGCAGC 300
QY      347  TGGGATCCAGGCCCCCTCTCTAGTCGGGAGCAAGCCAGGCTTGCACTTTATAGACCA 406
      301  TGGGATCCAGGCCCCCTCTCTAGTCGGGAGCAAGCCAGGCTTGCACTTTATAGACCA 360
QY      407  CCGGGCTGGCTTATCGGAGGGTCAAAAGTTGAGTGGCTGTGGATGCTCTGTACGG 466
      361  CCGGGCTGGCTTATCGGAGGGTCAAAAGTTGAGTGGCTGTGGATGCTCTGTACGG 420
QY      467  GAAGTCTCTGAGGATGAGCAGTACAGGAGTGGCGGCGGAGCCCAACCCCAAGCAA 526
      421  GAAGTCTCTGAGGATGAGCAGTACAGGAGTGGCGGCGGAGCCCAACCCCAAGCAA 480
QY      527  GATGCGGAAGCTCTTTCAGTTTTCACACAGCCCTGGAACCTGGACCTGCAAGGACTTCTCT 586
      481  GATGCGGAAGCTCTTTCAGTTTTCACACAGCCCTGGAACCTGGACCTGCAAGGACTTCTCT 540
QY      587  CCAGGCCCCTAAGGAGTCCCAGTCTCTAGTCTCTAGTGGTGGAGACCTGGAGCGGAGCTGAGGCTC 646
      541  CCAGGCCCCTAAGGAGTCCCAGTCTCTAGTCTCTAGTGGTGGAGACCTGGAGCGGAGCTGAGGCTC 600
QY      647  CTTCCACAGACACTCCGGTCAGCCCTGGCAATCCCAACCAATCATCTGAATCTGATC 706
      601  --TTCCAGCACACTCCGGTCAGCCCTGGG--AATCCCAACCAATCATCTGAATCTGATC 657
QY      707  TTTTATACAAATATACGAAAAGC 731
      658  TTTTATACAAATATACGAAAAGC 682

RESULT 15
BM998354/c
LOCUS   BM998354
DEFINITION UI-H-DT1-awc-p-21-0-UI-s1 NCI_CGAP_DT1 Homo sapiens cDNA clone
ACCESSION IMAGE:5887748 3', mRNA sequence.
VERSION   BM998354
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 638)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
           Tumor Gene Index
           Unpublished
           Contact: Robert Strausberg, Ph.D.
           Email: cgabbs-r@mail.nih.gov
           Tissue Procurement: Dr. Jose Mercuende
           CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
           CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
           DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
           Clone Distribution: Clone distribution information can be found
           through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
           Seq primer: M13 FORWARD
           POLYA=yes.
           Location/Qualifiers
             1..638
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:5887748"
               /tissue_type="Metastatic Chondrosarcoma"
               /dev_stage="Adult"
               /lab_host="DH10B (Life Technologies)"
               /clone_lib="NCI CGAP DT1"
               /note="Organ: Lung, Vector: pTT3-Pac (Pharmacia) with a
               modified polylinker; Site_1: EcoR I; Site_2: Not I;
               NCI CGAP DT1 is a normalized cDNA library containing the
               following tissue(s): Metastatic Chondrosarcoma in Lung.
               The library was constructed according to Bonaldo, Lennon

```

and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pTV3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AACTGTTCCG.

TAG LIB=UI-H-Df1
TAG_TISSUE=lung metastatic chondrosarcoma
TAG_SEQ=AACTGTTCCG"

BASE COUNT 108 a 185 c 201 g 144 t
ORIGIN

Query Match		83.4%;	Score 617.2;	DB 12;	Length 638;
Best Local Similarity		99.5%;	Pred. No. 1.3e-128;		
Matches 619;		Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
QY	119	GTTCAGAGCTGAAGCTGTCGGTCCGCTGCGCGAGGGCTACGGGCGCATCCCGCGGG	178		
Db	638	GTCAAGAGCTGAAGCTGCTGCGTCCGCTGCGGAGGGCTACGGGCGCATCCCGCGGG	579		
QY	179	CGCGCTGCTGTCCATGGACGCTTGACCTACCGACAAGCTGGTCAGCTTCTACCTGGA	238		
Db	578	CGCGCTGCTGTCCATGGACGCTTGACCTACCGACAAGCTGGTCAGCTTCTACCTGGA	519		
QY	239	GACCTACGGCGCGGAGCTACCGCTAACGTGCTGCCGACATGGGCTGCAGGAGATGGC	298		
Db	518	GACCTACGGCGCGGAGCTACCGCTAACGTGCTGCCGACATGGGCTGCAGGAGATGGC	459		
QY	299	CGGGCAGCTGCAGGCGGCGCACCGACAGGGCTCTGGAGCCGCCAGCTGGGATCCAGGC	358		
Db	458	CTGGCAGCTGCAGGCGGCGCACCGACAGGGCTCTGGAGCCGCCAGCTGGGATCCAGGC	399		
QY	359	CCCTCCTCAGTCCGSCAGCCAAAGCAGGCTGCACCTTTATAGACCAGCACCGGGCTGCGCT	418		
Db	398	CCCTCCTCAGTCCGSCAGCCAAAGCAGGCTGCACCTTTATAGACCAGCACCGGGCTGCGCT	339		
QY	419	TATCGGAGGGTCACAAAGTTGAGTGGCTGTGGATGCTCTGTACGGGAAGGTCTCTGAC	478		
Db	338	TATCGGAGGGTCACAAAGTTGAGTGGCTGTGGATGCTCTGTACGGGAAGGTCTCTGAC	279		
QY	479	GGATGAGCAGTACACGCGAGTGGGGCCGAGCCACCAACCAAGCAAGATCGGAAGCT	538		
Db	278	GGATGAGCAGTACACGCGAGTGGGGCCGAGCCACCAACCAAGCAAGATCGGAAGCT	219		
QY	539	CTTCAGTTTCACACAGCTGGAACTGGACCTGCAAGGACTTGTCTCTCCAGGCCCTAAG	598		
Db	218	CTTCAGTTTCACACAGCTGGAACTGGACCTGCAAGGACTTGTCTCTCCAGGCCCTAAG	159		
QY	599	GGAGTCCCAAGTCTTACCTGGTGGAGACCTGGAGCGGAGCTGAGGCTCTTCCAGCAAC	658		
Db	158	GGAGTCCCAAGTCTTACCTGGTGGAGACCTGGAGCGGAGCTGAGGCTCTTCCAGCAAC	99		
QY	659	ACTCCGGTCCAGCCCTGGCAATCCCAACCAATCATCTGAATCTGATCTTTTATACACA	718		
Db	98	ACTCCGGTCCAGCCCTGGCAATCCCAACCAATCATCTGAATCTGATCTTTTATACACA	39		
QY	719	ATATACGAAAAGCAGCTTGAA	740		
Db	38	ATATACGAAAAGCAGCTTGAA	17		

Search completed: January 29, 2004, 13:05:55
Job time : 1480.73 secs

Result No.	Score	Match	Query	Length	DB	ID	Description	§
1	740	100.0	740	4	US-09-340-620A-48	Sequence 48, Appl		
2	585	79.1	585	4	US-09-340-620A-50	Sequence 50, Appl		
3	336	45.4	777	4	US-09-340-620A-60	Sequence 60, Appl		
4	316.2	42.7	579	4	US-09-340-620A-62	Sequence 62, Appl		
5	53	7.2	909	4	US-09-352-991A-3958	Sequence 2988, Ap		
6	53	7.2	1218	4	US-09-352-991A-3146	Sequence 3146, Ap		
C	7	53	7.2	1380	4	US-09-252-991A-2822	Sequence 2822, Ap	
	8	53	7.2	1500	4	US-09-252-991A-3050	Sequence 3050, Ap	
	9	50.6	6.8	2079	4	US-09-252-991A-3097	Sequence 3097, Ap	
10	49.8	6.7	33529	3	US-09-144-085-3	Sequence 3, Appl		
C	11	49.2	6.6	50937	3	US-09-428-517-1	Sequence 1, Appl	
	12	49	6.6	2610	4	US-09-352-991A-1107	Sequence 1107, Ap	
	13	49	6.6	3114	4	US-09-252-991A-977	Sequence 977, App	
14	49	6.6	3195	4	US-09-252-991A-937	Sequence 937, App		
C	15	48.8	6.6	468	4	US-09-252-991A-10961	Sequence 10961, A	
	16	48.8	6.6	795	4	US-09-252-991A-11104	Sequence 11104, A	
	17	48.8	6.6	1044	4	US-09-252-991A-11159	Sequence 11159, A	
C	18	48.8	6.6	1557	4	US-09-252-991A-11093	Sequence 11093, A	
	19	48	6.5	30001	1	US-08-125-468-1	Sequence 1, Appl	
	20	48	6.5	30001	2	US-08-474-933-1	Sequence 1, Appl	
C	21	47.8	6.5	558	4	US-09-252-991A-6608	Sequence 6608, Ap	
	22	47.8	6.5	645	4	US-09-252-991A-6723	Sequence 6723, Ap	
	23	47.8	6.5	984	4	US-09-252-991A-6664	Sequence 6664, Ap	
24	46.6	6.3	800	5	PCT-US95-04801-4	Sequence 4, Appl		
25	46.2	6.2	23673	4	US-09-773-816-1	Sequence 1, Appl		
26	45	6.1	876	4	US-09-252-991A-12184	Sequence 12184, A		
27	45	6.1	1392	4	US-09-252-991A-12249	Sequence 12249, A		

Db 241 CCTACGGCGCCGAGCTCACCGCTAAAGTGTGCGGACATGGGCTTCAGGAGATGGCG 300
 Qy 301 GCGAGCTGAGGGCGGCGACCGACCGAGGCTTCGGAGCGGCGGAGTGGATCCAGGCGC 360
 Db 301 GCGAGCTGAGGGCGGCGACCGACCGAGGCTTCGGAGCGGCGGAGTGGATCCAGGCGC 360
 Qy 361 CTCTCAGTGGGAGCGCAAGCGAGGCTGCACCTTTATAGACAGCACCGGGCTGCGCTTA 420
 Db 361 CTCTCAGTGGGAGCGCAAGCGAGGCTGCACCTTTATAGACAGCACCGGGCTGCGCTTA 420
 Qy 421 TCGGAGGCTCAAAACGTTGAGTGGCTGCTGATGCTCTGTACGGAAAGTCTTGAACGG 480
 Db 421 TCGGAGGCTCAAAACGTTGAGTGGCTGCTGATGCTCTGTACGGAAAGTCTTGAACGG 480
 Qy 481 ATGAGCAGTACAGGAGGCTGCGGGCGGAGCCACCAACCAAGCAGATGCGGAAGCTCT 540
 Db 481 ATGAGCAGTACAGGAGGCTGCGGGCGGAGCCACCAACCAAGCAGATGCGGAAGCTCT 540
 Qy 541 TCAGTTTACACAGGAGGCTGGAACCTGCAAGGACTTGTCTCCAGGCGCTTAAGGG 600
 Db 541 TCAGTTTACACAGGAGGCTGGAACCTGCAAGGACTTGTCTCCAGGCGCTTAAGGG 600
 Qy 601 AGTCCAGTCTTACCTGTTGGAGGACCTGGAGCGGAGCTGAGGCTCTTCCAGCAACAC 660
 Db 601 AGTCCAGTCTTACCTGTTGGAGGACCTGGAGCGGAGCTGAGGCTCTTCCAGCAACAC 660
 Qy 661 TCGGTCAGGCGCTGCAATCCCAACCAATCATCTGATCTGATCTTTTATACAAAT 720
 Db 661 TCGGTCAGGCGCTGCAATCCCAACCAATCATCTGATCTGATCTTTTATACAAAT 720
 Qy 721 ATACGAAAAGCCAGCTTGAA 740
 Db 721 ATACGAAAAGCCAGCTTGAA 740

RESULT 2
 US-09-340-620A-50
 ; Sequence 50, Application US/09340620A
 ; Patent No. 6482933
 ; GENERAL INFORMATION:
 ; APPLICANT: Bertin, John
 ; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
 ; FILE REFERENCE: 07334-124001
 ; CURRENT APPLICATION NUMBER: US/09/340,620A
 ; CURRENT FILING DATE: 1999-06-28
 ; PRIOR APPLICATION NUMBER: US/09/245,281
 ; PRIOR FILING DATE: 1999-02-05
 ; PRIOR APPLICATION NUMBER: US/09/207,359
 ; PRIOR FILING DATE: 1998-12-08
 ; PRIOR APPLICATION NUMBER: US/09/099,041
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: US/09/019,942
 ; PRIOR FILING DATE: 1998-02-06
 ; NUMBER OF SEQ ID NOS: 71
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 50
 ; LENGTH: 585
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-340-620A-50

Query Match 79.1%; Score 585; DB 4; Length 585;
 Best Local Similarity 100.0%; Pred. No. 8.5e-123;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 54 ATGGGCGCGCGCGAGCGCATCTCTGGATGGCTGGAGAACTGACCGCGGAGGAGCTC 113
 Db 1 ATGGGCGCGCGCGAGCGCATCTCTGGATGGCTGGAGAACTGACCGCGGAGGAGCTC 60
 Qy 114 AAGAGTTCAAGCTGAAGTGTGTGTCGTCGCGGAGGCTTACGGGCGCATCCCG 173
 Db 61 AAGAGTTCAAGCTGAAGTGTGTGTCGTCGCGGAGGCTTACGGGCGCATCCCG 120

Qy 174 CCGGCGCGGCTGCTCTCATGAGCGCTTGGACCTCAACGACCAAGCTGGTCTCAGCTTCTAC 233
 Db 121 CCGGCGCGGCTGCTCTCATGAGCGCTTGGACCTCAACGACCAAGCTGGTCTCAGCTTCTAC 180
 Qy 234 CTGGAGAGCTTACGGCGCGAGCTCAACGCTAAAGTGTGGCGGACATGGGCTTCAGGAG 293
 Db 181 CTGGAGAGCTTACGGCGCGAGCTCAACGCTAAAGTGTGGCGGACATGGGCTTCAGGAG 240
 Qy 294 ATGGCGCGGAGCTGCGAGCGGCGCACGACCAAGGCTCTGGAGCGCGGAGCTGGGATC 353
 Db 241 ATGGCGCGGAGCTGCGAGCGGCGCACGACCAAGGCTCTGGAGCGCGGAGCTGGGATC 300
 Qy 354 CAGGCGCTCTCTCAGTCTGGGAGCAAGCGAGCGCTTGTATAGACAGCACCGGCT 413
 Db 301 CAGGCGCTCTCTCAGTCTGGGAGCAAGCGAGCGCTTGTATAGACAGCACCGGCT 360
 Qy 414 GCGCTTATCGGAGGCTCAAAAAGTGTGAGTGGCTGTGGATGCTCTGTACGGGAAGTCT 473
 Db 361 GCGCTTATCGGAGGCTCAAAAAGTGTGAGTGGCTGTGGATGCTCTGTACGGGAAGTCT 420
 Qy 474 CTGACGGATGAGCAGTACGAGCAGTGGGCGGAGCCCAACCAAGCAAGATGCGG 533
 Db 421 CTGACGGATGAGCAGTACGAGCAGTGGGCGGAGCCCAACCAAGCAAGATGCGG 480
 Qy 534 AAGCTCTTCTCAGTTTACACACCGCTTGAACCTGGACCTGCAAGGACTTGTCTCTCAGGCT 593
 Db 481 AAGCTCTTCTCAGTTTACACACCGCTTGAACCTGGACCTGCAAGGACTTGTCTCTCAGGCT 540
 Qy 594 CTAAGGAGTCCAGTCTCTACCTGGTGGAGGAGCTTGGAGCGGAGC 638
 Db 541 CTAAGGAGTCCAGTCTCTACCTGGTGGAGGAGCTTGGAGCGGAGC 585

RESULT 3
 US-09-340-620A-60
 ; Sequence 60, Application US/09340620A
 ; Patent No. 6482933
 ; GENERAL INFORMATION:
 ; APPLICANT: Bertin, John
 ; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
 ; FILE REFERENCE: 07334-124001
 ; CURRENT APPLICATION NUMBER: US/09/340,620A
 ; CURRENT FILING DATE: 1999-06-28
 ; PRIOR APPLICATION NUMBER: US/09/245,281
 ; PRIOR FILING DATE: 1999-02-05
 ; PRIOR APPLICATION NUMBER: US/09/207,359
 ; PRIOR FILING DATE: 1998-12-08
 ; PRIOR APPLICATION NUMBER: US/09/099,041
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: US/09/019,942
 ; PRIOR FILING DATE: 1998-02-06
 ; NUMBER OF SEQ ID NOS: 71
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 60
 ; LENGTH: 777
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (89)...(667)
 US-09-340-620A-60

Query Match 45.4%; Score 336; DB 4; Length 777;
 Best Local Similarity 72.9%; Pred. No. 7.4e-67;
 Matches 461; Conservative 0; Mismatches 165; Indels 6; Gaps 2;
 Qy 18 GGGTCAGCGGCGGAGCGCGGAGCTCTGGAGCATGGGCGCGCGCGAGCCATC 77
 Db 53 GAGTAAAGGTGACCGCGGCTGCCACCCAGAGCCATGGGCGGCGAGATGCCATC 112
 Qy 78 CTGGATGCGCTGGAGAACTTGAACCGCGGAGGAGCTCAAGAGTTCAAGCTGAAGTCTGTG 137

Db 113 CTGGAGCGCTCTTGAACACTTGTACGGGATGAACCTCAAAAAGTTCAAGATGAAGCTGCTG 172
Qy 138 TCGGTGCCCTCGCGAGGGCTACGGGCGCATCCCGCGGGCGCGCTGCTGTCCATGGAC 197
Db 173 ACAGTGCAACTCGGAGAGGCTATGGGCGCATCCACGCGGGGCCCTGCTGCAGATGGAC 232
Qy 198 GCGTTGGACCTCACCGACAAGCTGTGCTTACCTGGAGACCTACCGGCGCGAGCTC 257
Db 233 GCATAGATCTCACTGACAACTTGTCACTACTATCTGGAGTCTGATGGCTTGGAGCTC 292
Qy 258 ACCGCTAAGCTCTGCGGCACATGGCCCTCGAGGAGATGGCGGCGAGCTGAGGGGCC 317
Db 293 ACAATGACTGTGCTTAGAGACATGGGCTTACAGGAGCTGGTGAAGCTGCAAAACG--- 349
Qy 318 AGCCACCAAGGCTCTGGAGCGCGCCAGCTGGGATCGAGGCCCTCTCAGTCGGCAGCC 377
Db 350 ACTAAGAAGAGCTGTGGAGCTGTGGCAGCTGAGCCAGTGTCTCTGCTCAGAGTACAGCC 409
Qy 378 AAGCCAGGCTCGACTTTATAGACCAAGCACCGGGCTGCGCTTATCGCGAGGCTCAAAAC 437
Db 410 AGAACAGG---ACACTTTGTGACCAAGCACAGCAAGCACTATTGCCAGGGTCAAGAA 466
Qy 438 GTTGAGTGTCTCTGATCTCTGTACCGGAAGTCTTACCGGATGAGCAGTACAGGCA 497
Db 467 GTGGACGGAGTGTGATCTTGTGATGCGAGTGTGCTGACTGAAGGACAGTACAGGCA 526
Qy 498 GTGGCGGCGAGCCACCAACCAAGCAAGATGCGGAAGCTCTTCAAGTTTACACAGCC 557
Db 527 GTTGTGAGAGACCAACCAAGCAAGATGAGGAAGCTTTCAGCTTTGTTCCATCC 586
Qy 558 TGGAACTGACCTGCAAGGACTTGTCTTCCAGGCCCTAAGGAGTCCAGTCTTACCTG 617
Db 587 TGGAACTGACCTGCAAGGACTTCTTCCAGGCCCTTCCAGGCCCTTGAAGAAATACATCCCTACTTG 646
Qy 618 GTGGAGGAGCTGAGCGGAGCTGAGGCTCCTT 649
Db 647 GTGATGGACCTGGAGCAGAGCTGAGGTATCTT 678

RESULT 4
US-09-340-620A-62
; Sequence 62, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340, 620A
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 62
; LENGTH: 579
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-340-620A-62

Query Match 42.7%; Score 316.2; DB 4; Length 579;
Best Local Similarity 73.7%; Pred. No. 1.9e-62;
Matches 431; Conservative 0; Mismatches 148; Indels 6; Gaps 2;
Qy 54 ATGGGCGGCGCGGCGGAGCCATCTCTGGATGCGCTGGAGAACCTGACCGCGGAGGAGCTC 113
Db 1 ATGGGCGGCGCGGAGATGCCATCTCTGGAGCTCTTGAAGAACTTGTGAGGGATGAATC 60

Qy 114 AAGAAAGTTCAAGCTGAAGCTGCTGTGCTGCGCTGCGGAGGCTACGGGCGCATCCCG 173
Db 61 AAAAAGTTCAAGATGAAGCTGCTGACAGTGAACCTGCGAGAAGGCTATGGGCGCATCCA 120
Qy 174 CGGGCGGCGCTGCTGCTCCATGAGCGCTTGGACCTACCGCAAGCTGCTGAGCTTCTAC 233
Db 121 CCGGGGCGCTGCTGAGATGAGCGCATAGATCTCATCTGCAAACTTGTGAGCTACTAT 180
Qy 234 CTGAGAGACTTACGGCGCGAGCTCACCGCTAACCTGTGCTGCGGACATGGGCTTGCAGGAG 293
Db 181 CTGAGTCTGATGGCTTGGAGCTCAATGACTGTGCTTAGAGACATGGCTTACAGGAG 240
Qy 294 ATGGCGGCGAGCTGCGAGGCGCCACCGACCAAGGCTCTGAGCGCGCGCAGCTGGGATC 353
Db 241 CTGGCTGAGCAGCTGCAAAACG---ACTAAAGAAAGTCTGGAGCTGTGGCAGCTGCAGCC 297
Qy 354 CAGGCGCTCTCTCAGTCTGGCAGCCAAAGCAGCGCTTATAGACAGCAGCAGGCT 413
Db 298 AGTGTCCCTGCTCAGAGTACAGCCAGAACAGG---ACATTTGTGGACAGCAGGCA 354
Qy 414 GCGCTTATCGGAGGCTCACAAACGTTGAGTGGCTGTGCTGTGATGCTCTGTACGGGAGGTC 473
Db 355 GCACCTCATTTGCAGGGTCAAGAACTGGACGAGTGTGCTGTGATGCTTGCATGCGAGTGTG 414
Qy 474 CTGACGGATGAGCAGTACAGGCACTGCGGGCGGAGCCACCAACCCAAAGCAAGATGCGG 533
Db 415 CTGACTGAAGGACAGTACAGGCACTTCTGTGAGAGACCCAGCCAGCAAGATGAGG 474
Qy 534 AGCTCTTTCAGTTTACACAGGCTTGGAGCTGAGCTGCAAGGACTTGTCTCTCCAGGCC 593
Db 475 AAGCTCTTTCAGTTTGTTCCTGGAACCTGACTTGAAGGACTTCTCTCTCCAGGCC 534
Qy 594 CTAAGGAGTCCAGCTCTCTGCTGAGGAGGCTTGGAGCGGAGG 638
Db 535 TTGAAGGAATACATCCCTACTTGTGATGAGCTTGGAGCAGG 579

RESULT 5
US-09-252-991A-2958
; Sequence 2958, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107195.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2958
; LENGTH: 909
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2958

Query Match 7.2%; Score 53; DB 4; Length 909;
Best Local Similarity 49.1%; Pred. No. 0.0027;
Matches 140; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

Qy 63 GCGCGGAGCGCCATCTGTGATGCGCTGGAGAACCTGACCGCGGAGGCTCAGAGTTTC 122
Db 576 GCGCGGCTCTCACCACCGAATACGAAGCGCGCAAGCGGTCAAGGAATCGCGGCTTC 635
Qy 123 AAGCTGAAGCTCTGTGCTGCGCTGCGCGAGGCTACGGCGCATCCCGGGGCGCG 182
Db 636 CGCTCTTCTGCTGTTTGGAGCAGAGCTTCCAGACTACGTACGTCTCTCCAGGCAAC 695
Qy 183 CTGCTGTCTCAGAGGCTTGGACCTTACCGCAAGAGCTGTCAGCTTCTTACCTGGAGCC 242

Db 696 GACCGCTACAGAACGCCCTGGACAGCGCGCGCAACCCCGAGCGCTTCATGAGAACTG 755
Qy 243 TAGCGCGCGAGCTACCCGCTACGCTGCTGCGGACATGGGCTGCGAGAGATGCCGG 302
Db 756 CAGCGCGCGGCTACGCCACCGATCCGAGTAGTACGCCCGCAAGGTGGCGCAGATCGCCAGA 815
Qy 303 CAGCTGCGAGCGCGCACGACGACGAGGCTCTGGAGCGCGCGCCAGCT 347
Db 816 CAGATGCAGACTACGAGCGCGCTGCGCGCGCGCGGACGCGCGCT 860

RESULT 6

US-09-252-991A-3146
; Sequence 3146, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 3146
; LENGTH: 1218
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3146

Query Match 7.2%; Score 53; DB 4; Length 1218;
Best Local Similarity 49.1%; Pred. No. 0.0029;
Matches 140; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

Qy 63 GCGCGCGAGCCATCTGTGATGCGTGGAGAACCTGACCGCGGAGGCTCAAGAATTC 122
Db 925 GCGCGCGCTCTACCAACCGAATACGAAGCGGCAAGCGGTCAAGGAAGTCCGCGGTTTC 984
Qy 123 AAGCTGAAGCTGTGTGCTGCGTGGAGAACCTGACCGCGGAGGCTACCGCGCGCGCG 182
Db 985 CGCTCTACTCTGCTGTGCTGAGCAGAGCTTCACGACTAGTCACTTCTTCTCAGGGCAAC 1044
Qy 183 CTGCTGTTCATGAGACCTTGGACCTCACCGAAGCTGGTCACTTCTTCTTCTGAGACC 242
Db 1045 GACCGCTACCAAGACCGCTTGGACAGCGCGCGCAACCCCGAGCGCTTCATGAGGAATG 1104
Qy 243 TAGCGCGCGAGCTACCGCTACGCTGCTGCGGACATGGGCTGCGAGAGATGCCGG 302
Db 1105 CAGCGCGCGGTACTACGCCACCGATCCGCAAGTACGCGCGCGCGCAAGGTGGCGCAGATCGCCAGA 1164
Qy 303 CAGCTGCGAGCGCGCACGACGACGAGGCTCTGGAGCGCGCGCCAGCT 347
Db 1165 CAGATGCAGACTACGAGCGCGCTGCGCGCGCGCGGACGCGCGCT 1209

RESULT 7

US-09-252-991A-2822/c
; Sequence 2822, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2822
; LENGTH: 1380
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2822

Query Match 7.2%; Score 53; DB 4; Length 1380;
Best Local Similarity 49.1%; Pred. No. 0.003;
Matches 140; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

Qy 63 GCGCGCGAGCCATCTGTGATGCGTGGAGAACCTGACCGCGGAGGCTCAAGAATTC 122
Db 321 GCGCGCGCTCTACCAACCGAATACGAAGCGGCAAGCGGTCAAGGAAGTCCGCGGTTTC 262
Qy 123 AAGCTGAAGCTGTGTGCTGCGTGGAGAACCTGACCGCGGAGGCTACCGCGGCGCG 182
Db 261 CGCTCTACTCTGCTGCTGAGCAGAGCTTCACGACTACGTCACTTCTTCTCAGGGCAAC 202
Qy 183 CTGCTGTCCATGGAGCGCTTGGACCTCACCGACAAAGCTGGTCACTTCTTCTGAGACC 242
Db 201 GACCGCTACGAGAACCGCTTGGACAGCGCCCGCAACCCCGAGCGCTTCATGAGGAATG 142
Qy 243 TAGCGCGCGAGCTCACCGCTTACGCTGCTGCGCGACATGGGCTTCAGAGATGCCGG 302
Db 141 CAGCGCGCGGTAGCGCCACCGATCCGCAAGTACGCGCGCAAGGTGGCGCAGATCGCCAGA 82
Qy 303 CAGCTGCGAGCGCGCACGACGACGAGGCTCTGGAGCGCGCGCCAGCT 347
Db 81 CAGATGCAGACTACGAGCGCGCTGCGCGCGCGCGGACGCGCGCT 37

RESULT 8

US-09-252-991A-3050
; Sequence 3050, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 3050
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3050

Query Match 7.2%; Score 53; DB 4; Length 1500;
Best Local Similarity 49.1%; Pred. No. 0.003;
Matches 140; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

Qy 63 GCGCGCGAGCCATCTGTGATGCGTGGAGAACCTGACCGCGGAGGCTCAAGAATTC 122
Db 920 GCGCGCGCTCTACCAACCGAATACGAAGCGGCAAGCGGTCAAGGAAGTCCGCGGTTTC 979
Qy 123 AAGCTGAAGCTGTGTGCTGCGTGGAGAACCTGACCGCGGAGGCTACCGCGGCGCG 182
Db 980 CGCTCTACTCTGCTGCTGAGCAGAGCTTCACGACTACGTCACTTCTTCTCAGGGCAAC 1039
Qy 183 CTGCTGTCCATGGAGCGCTTGGACCTCACCGACAAAGCTGGTCACTTCTTCTGAGACC 242
Db 1040 GACCGCTACGAGAACCGCTTGGACAGCGCCCGCAACCCCGAGCGCTTCATGAGGAATG 1099
Qy 243 TAGCGCGCGAGCTCACCGCTTACGCTGCTGCGCGACATGGGCTTCAGAGATGCCGG 302
Db 1100 CAGCGCGCGGTAGCGCCACCGATCCGCAAGTACGCGCGCGCGCAAGGTGGCGCAGATCGCCAGA 1159

QY 326 GGCTCTGGAGCGCGCCAGCTGGGATCCAGGCCCCCTCTCAGTCGGCAGCCAAAGCCAGG 385
 Db 1510 TCGGCCCGCTACCGCTGCTCGGTACGGGGCCGACTCTCAGCGCGCGCGCGCGCG 1451
 QY 386 CC 387
 Db 1450 CC 1449
 RESULT 12
 US-09-252-991A-1107/c
 ; Sequence 977, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 1107
 ; LENGTH: 2610
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-1107

Query Match 6.6%; Score 49; DB 4; Length 2610;
 Best Local Similarity 49.8%; Pred. No. 0.027;
 Matches 124; Conservative 0; Mismatches 125; Indels 0; Gaps 0;
 QY 62 CGCGCGCGAGCCATCTCGATCGCTGGAGAACCTGACCGCGGAGAGCTCAAGAAGTT 121
 Db 1942 CGAGCGCGCATTCCTCGCTGCACGCGCCGAGGAGTGAACCATCCGCTGTGCTGGAGCG 1883
 QY 122 CAAGCTGAAGCTGCTCTCGTGGTCCGCTGCGGAGGCTACGGCGCATCCCGGGGGCGC 181
 Db 1882 CATCCGCCAACTGCGCGCGGAGCTTCCTGTTCTTCTTACTACCGCGCGCTGCTCGGCGC 1823
 QY 182 GCTGCTGCTCATGAGCGCTTGAGCTCCTGACCTACCGAGAGCTGGTCACTTCTACTCGAGAC 241
 Db 1822 CGAGCTGCTGCTGCGCGCGGAGGCGCTTACACCTGACAGGTTCGCTGCTGCGCGG 1763
 QY 242 CTACGCGCGCGAGCTCACCGCTAAACGTGCTGCGCGACATGGGCGCTGACGAGATGCGCGG 301
 Db 1762 CTACGCGCGAGAGCGCGCGGAACTGGGTGCTGCTCAACGCGGAAACGACAGACCGGGGT 1703
 QY 302 GCAGCTGCA 310
 Db 1702 GACCTTGCA 1694
 RESULT 13
 US-09-252-991A-977
 ; Sequence 977, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 977

; LENGTH: 3114
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-977
 Query Match 6.6%; Score 49; DB 4; Length 3114;
 Best Local Similarity 49.8%; Pred. No. 0.028;
 Matches 124; Conservative 0; Mismatches 125; Indels 0; Gaps 0;
 QY 62 CGCGCGCGAGCCATCTCGATCGCTGGAGAACCTGACCGCGGAGAGCTCAAGAAGTT 121
 Db 1281 CGAGCGCGCATTCCTCGCTGCACGCGCCGAGGACGTGAACCATCCGCTGTGCTGGAGCG 1340
 QY 122 CAAGCTGAAGCTGCTGCTCGCTGCGCGAGGGCTACGCGGCGCATCCCGCGGGGGCGC 181
 Db 1341 CATCGCCAACTGCGCGCGGACTTCTGTTCTTCTTACTACCGCGCGCTGCTCGGCGC 1400
 QY 182 GCTGCTGCTCATGAGCGCTTGGACCTCACGACAAAGCTGTGCTAGCTTCTTACTCGGAGAC 241
 Db 1401 CGAGCTGCTGCTGCGCGCGGCGCATCAAACTGCACGCTTCTGCTGCTGCTGCTGCGCG 1460
 QY 242 CTACGCGCGCGAGCTCACCGCTAACGCTGCTGCGGACATGGGCGCTGACGAGATGCGCGG 301
 Db 1461 CTACGCGGAGAGCGCGCGGAACTGGGTGCTGCTCAACGCGGAAACGACAGACCGGGGT 1520
 QY 302 GCAGCTGCA 310
 Db 1521 GACCTTGCA 1529
 RESULT 14
 US-09-252-991A-937
 ; Sequence 937, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 937
 ; LENGTH: 3195
 ; TYPE: DNA
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-937

Query Match 6.6%; Score 49; DB 4; Length 3195;
 Best Local Similarity 49.8%; Pred. No. 0.028;
 Matches 124; Conservative 0; Mismatches 125; Indels 0; Gaps 0;
 QY 62 CGCGCGCGAGCCATCTCGATCGCTGGAGAACCTGACCGCGGAGAGCTCAAGAAGTT 121
 Db 478 CGAGCGCGCATTCCTCGCTGCACGCGCCGAGGACGTGAACCATCCGCTGTGCTGGAGCG 537
 QY 122 CAAGCTGAAGCTGCTGCTCGCTGCGCGAGGGCTACGCGGCGCATCCCGCGGGGGCGC 181
 Db 538 CATCGCCAACTGCGCGCGGACTTCTGTTCTTCTTACTACCGCGCGCTGCTCGGCGC 597
 QY 182 GCTGCTGCTCATGAGCGCTTGGACCTCACGACAAAGCTGTGCTAGCTTCTTACTCGGAGAC 241
 Db 598 CGAGCTGCTGCTGCGCGCGGCGCATCAAACTGCACGCTTCTGCTGCTGCTGCTGCGCG 657
 QY 242 CTACGCGCGCGAGCTCACCGCTAACGCTGCTGCGGACATGGGCGCTGACGAGATGCGCGG 301
 Db 658 CTACGCGGAGAGCGCGCGGAACTGGGTGCTGCTCAACGCGGAAACGACAGACCGGGGT 717
 QY 302 GCAGCTGCA 310

Search completed: January 29, 2004, 13:11:29
Job time : 41.6087 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2004, 03:03:41 ; Search time 235.019 Seconds
(without alignments)
11475.528 Million cell updates/sec

Title: US-09-996-617-7
Perfect score: 740
Sequence: 1 cgcgtccggctgcagcgggg.....atcacgaagcagcttgaa 740

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2434939 seqs, 1822278265 residues

Total number of hits satisfying chosen parameters: 4869878

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	740	100.0	740	9	US-09-728-721-48
2	740	100.0	740	10	US-09-996-617-7
3	740	100.0	740	10	US-09-996-617-10
4	740	100.0	740	10	US-09-841-879B-4
5	740	100.0	740	10	US-09-841-879B-19
6	740	100.0	740	15	US-10-295-981-48
7	733	99.1	745	12	US-10-240-145-12
8	733	99.1	779	12	US-10-131-410-21
9	721.2	97.5	806	15	US-10-106-698-1144
10	697.6	94.3	811	9	US-09-925-301-278
11	585	79.1	585	9	US-09-728-721-50
12	585	79.1	585	10	US-09-996-617-9
13	585	79.1	585	10	US-09-841-879B-6
14	585	79.1	585	15	US-10-295-981-50
15	531	71.8	639	15	US-10-106-698-1145

16	384.6	52.0	432	9	US-09-925-299-648	Sequence 648, App	
17	384.6	52.0	432	11	US-09-925-299-648	Sequence 648, App	
18	336	45.4	777	9	US-09-728-721-60	Sequence 60, Appl	
19	336	45.4	777	10	US-09-841-879B-1	Sequence 1, Appl	
C	20	336	45.4	777	10	US-09-841-879B-18	Sequence 18, Appl
21	336	45.4	777	15	US-10-295-981-60	Sequence 60, Appl	
22	316.2	42.7	579	9	US-09-728-721-62	Sequence 62, Appl	
23	316.2	42.7	579	15	US-10-295-981-62	Sequence 62, Appl	
C	24	309.8	41.9	321	10	US-09-880-107-1066	Sequence 1066, Ap
25	293.8	39.7	579	10	US-09-841-879B-3	Sequence 3, Appl	
C	26	244.6	33.1	551	12	US-10-191-803-602	Sequence 602, App
27	169.4	22.9	270	11	US-09-965-621-27	Sequence 27, Appl	
28	169.4	22.9	270	12	US-10-407-866-27	Sequence 27, Appl	
29	151.2	20.4	437	12	US-10-062-674-985	Sequence 985, App	
30	146.4	19.8	174	12	US-10-242-535A-45706	Sequence 45706, A	
31	115	15.5	4200	10	US-09-388-221-3	Sequence 3, Appl	
32	115	15.5	4287	10	US-09-996-617-5	Sequence 5, Appl	
33	115	15.5	4287	10	US-09-931-071-5	Sequence 5, Appl	
34	115	15.5	4332	10	US-09-388-221-5	Sequence 5, Appl	
35	115	15.5	4422	10	US-09-388-221-1	Sequence 1, Appl	
36	115	15.5	5100	11	US-09-956-712-12	Sequence 12, Appl	
37	115	15.5	5444	10	US-09-996-617-1	Sequence 1, Appl	
38	115	15.5	5444	10	US-09-931-071-1	Sequence 1, Appl	
39	115	15.5	5444	11	US-09-956-712-3	Sequence 3, Appl	
40	115	15.5	6531	11	US-09-956-712-11	Sequence 11, Appl	
41	111.4	15.1	2657	11	US-09-895-298-22	Sequence 22, Appl	
42	110.6	14.9	96649	11	US-09-956-712-10	Sequence 10, Appl	
43	98.6	13.3	180	12	US-10-240-145-98	Sequence 98, Appl	
C	44	74	394	9	US-09-864-761-21179	Sequence 21179, A	
45	70.6	9.5	2708	12	US-10-407-866-65	Sequence 65, Appl	

ALIGNMENTS

RESULT 1
US-09-728-721-48
; Sequence 48, Application US/09728721
; Patent No. US20020061845A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREO
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/728,721
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)...(638)
US-09-728-721-48

Query Match 100.0%; Score 740; DB 9; Length 740;
Best Local Similarity 100.0%; Pred. No. 1.1e-196;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CGCGTCCGCGTGCAGCGGGTGCAGCGCGCGCGGATCCTGGAGCCATGGGC	60
Db	1	CGCGTCCGCGTGCAGCGGGTGCAGCGCGCGGATCCTGGAGCCATGGGC	60
QY	61	GCGCGCGACGCCATCCTGGATGCGTGAGAACCTGACCGCGGAGAGCTCAAGAGT	120


```

; PRIOR APPLICATION NUMBER: 09/428,252
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-996-617-10

```

Query Match	100.0%	Score	740	DB	10	Length	740		
Best Local Similarity	100.0%	Pred. No.	1.1e-196						
Matches	740	Conservative	0	Mismatches	0	Indels	0	Gaps	0

Qy	1	CGCGTCGGGCTGCAGCGGGGTACGCGCGGCAGCGCGCGGGATCTCTGGAGCCATGGGGC	60
Db	740	CGCGTCCGGCTGCAGCGGGGTACGCGCGGCAGCGCGCGGGATCTCTGGAGCCATGGGGC	681
Qy	61	GC GCGCGCGAGCCCATCTCTGGATCGCTTGGAGAACTTGACCGCCGAGGAGTCTCAAGAA	120
Db	680	GCGCGCGGAGCCCATCTCTGGATCGCGCTGGAGAACTTGACCGCCGAGGAGTCTCAAGAA	621
Qy	121	TCAAGCTGAAGTCTGCTCGTGGCGCTGCGCGAGGGCTACGGGGGCATCCGCGGGGGCG	180
Db	620	TCAAGCTGAAGTCTGCTCGGTGCGCTGCGCGAGGGCTACGGGGGCATCCGCGGGGGCG	561
Qy	181	CGCTGCTGTCCATGAGAGCCCTTGACCTCACCGCAAGCTGGTCAAGCTTCTACCTGGAGA	240
Db	560	CGCTGCTGTCCATGGAAGCCCTTGACCTCACCGCAAGCTGGTCAAGCTTCTACCTGGAGA	501
Qy	241	CTTACGCGCGGAGGTACCGCTAACGCTGTGCGGGAATGGGCTTGAGGAGATGGCGG	300
Db	500	CTTACGCGCGGAGGTACCGCTAACGCTGTGCGGGAATGGGCTTGAGGAGATGGCGG	441
Qy	301	GGCAGCTGCAGGCGGCGACACAGAGGCTCTGAGAGCGCGCCAGCTGGGATCCAGGGCCC	360
Db	440	GGCAGCTGCAGGCGGCGCACAGAGGCTCTGAGAGCGCGCCAGCTGGGATCCAGGGCCC	381
Qy	361	CTCTCTAGTCGGCAGCCAGCCAGGCTGCACTTTATAGACCAGCACCGGCTCGCGTTA	420
Db	380	CTCTCTAGTCGGCAGCCAGCCAGGCTGCACTTTATAGACCAGCACCGGCTCGCGTTA	321
Qy	421	TCGCGAGGGTCACAAACGTTGAGTGGCTGCTGGATGCTCTGTA CGGGAAGGTCCTGACGG	480
Db	320	TCGCGAGGGTCACAAACGTTGAGTGGCTGCTGGATGCTCTGTA CGGGAAGGTCCTGACGG	261
Qy	481	ATGAGCAGTATCAGGCGAGTGGCGGCGGAGCCACCAACCCCAAGCAAGATGCGGAAGCTCT	540
Db	260	ATGAGCAGTATCAGGCGAGTGGCGGCGGAGCCACCAACCCCAAGCAAGATGCGGAAGCTCT	201
Qy	541	TCAGTTTACACAGAGCTGGAACTGGACCTGGAAGGACTTGCTCTCTCAGGCGCCTTAAGG	600
Db	200	TCAGTTTACACAGAGCTGGAACTGGACCTGGAAGGACTTGCTCTCTCAGGCGCCTTAAGG	141
Qy	601	AGTCCCAGTCTCTACTGTTGGAGGACCTGGAGCGGAGCTGAGGCTCTCTCCACGACAC	660
Db	140	AGTCCCAGTCTCTACTGTTGGAGGACCTGGAGCGGAGCTGAGGCTCTCTCCACGACAC	81
Qy	661	TCCGGTCAGCCCCCTGGCAATCCCACCAATCATCTGAACTCTGATCTTTTTATACCAAT	720
Db	80	TCCGGTCAGCCCCCTGGCAATCCCACCAATCATCTGAACTCTGATCTTTTTATACCAAT	21
Qy	721	ATACGAAAGCCAGCTTGAA	740
Db	20	ATACGAAAGCCAGCTTGAA	1

RESULT 4
US-09-841-879B-4
; Sequence 4, Application US/09841879B
; Patent No. US20020142979A1

```

; GENERAL INFORMATION:
; APPLICANT: Bertin. John
;
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
;
; FILE REFERENCE: 07334-330001
;
; CURRENT APPLICATION NUMBER: US/09/841,879B
;
; CURRENT FILING DATE: 2001-04-24
;
; PRIOR APPLICATION NUMBER: US 09/728,721
;
; PRIOR FILING DATE: 2000-12-01
;
; PRIOR APPLICATION NUMBER: US 09/340,620
;
; PRIOR FILING DATE: 1999-06-28
;
; NUMBER OF SEQ ID NOS: 19
;
; SOFTWARE: FastSeq for Windows Version 4.0
;
; SEQ ID NO 4
;
; LENGTH: 740
;
; TYPE: DNA
;
; ORGANISM: Homo sapiens
;
; FEATURE:
;
; NAME/KEY: CDS
;
; LOCATION: (54) ... (638)
;
; US-09-841-879B-4

```

Query Match	100.0.0.%;	Score 740;	DB 10;	Length 740;
Best Local Similarity	100.0.0.%;	Pred. NO. 1.1e-196;		
Matches 740;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	CGCGTCCGGCTGCAGCGGGTGAGCGGCGCAGCGCGCGGGATCCCTGGAGCCATGGGCG	60	
Db	1	CGCGTCCGGCTGCAGCGGGTGAGCGGCGCAGCGCGCGGGATCCCTGGAGCCATGGGCG	60	
Qy	61	GGCGGCGCAGCGCATCCTCGATCGCTGCGAGAACCTTGACCGCCGAGGAGCTCAAGAAGT	120	
Db	61	GGCGGCGCAGCCATCCTGATCGCTGCGAGAACCTTGACCGCCGAGGAGCTCAAGAAGT	120	
Qy	121	TCAAGCTGAAGCTGCTGCTGGTCCGCTGCGGAGGGCTACGCGGCGGCGG	180	
Db	121	TCAAGCTGAAGCTGCTGCTGGTCCGCTGCGGAGGGCTACGCGGCGGCGG	180	
Qy	181	CGCTGCTGTCATGGAAGCCCTTGACCTCACCGACAAGCTGGTCAGCTTCTACCTGGAGA	240	
Db	181	CGCTGCTGTCATGGAAGCCCTTGACCTCACCGACAAGCTGGTCAGCTTCTACCTGGAGA	240	
Qy	241	CCTACGGCGCGAGCTCACCGCTAACGTGCTGCGCGACATGGGCGCTCGAGAGATGGCCG	300	
Db	241	CCTACGGCGCGAGCTCACCGCTAACGTGCTGCGCGACATGGGCGCTCGAGAGATGGCCG	300	
Qy	301	GGAGCTGCAGGGCGGCACGCGACAGGGCTCTGGAGCCGCGCCAGCTGGAGTCCAGGCC	360	
Db	301	GGAGCTGCAGGGCGGCACGCGACAGGGCTCTGGAGCCGCGCCAGCTGGAGTCCAGGCC	360	
Qy	361	CTCTCAGTCGGGAGCCAAAGCCAGGCGCTGCACCTTTATAGACCAGCCGGGCTCGCTT	420	
Db	361	CTCTCAGTCGGGAGCCAAAGCCAGGCGCTGCACCTTTATAGACCAGCCGGGCTCGCTT	420	
Qy	421	TCGCGAGGGTCAAAAAGTTGAGTGGCTGCTGATGCTCTGTAACGGAAGTCTCTGACGG	480	
Db	421	TCGCGAGGGTCAAAAAGTTGAGTGGCTGCTGATGCTCTGTAACGGAAGTCTCTGACGG	480	
Qy	481	ATGAGCAGTACAGGAGTGGGCGCCAGGCCACCAACCCAAAGATGCGGAGTCT	540	
Db	481	ATGAGCAGTACAGGAGTGGGCGCCAGGCCACCAACCCAAAGATGCGGAGTCT	540	
Qy	541	TCAGTTTCACACGAGCTGGAATCTGGACCTGCAGGACTTGTCTCTCCAGGCCCTAAGG	600	
Db	541	TCAGTTTCACACGAGCTGGAATCTGGACCTGCAGGACTTGTCTCTCCAGGCCCTAAGG	600	
Qy	601	AGTCCCAAGTCTTACCTGGTGGAGACCTGGAGCGGAGCTGAGGCTCTCTCCAGCAAC	660	
Db	601	AGTCCCAAGTCTTACCTGGTGGAGACCTGGAGCGGAGCTGAGGCTCTCTCCAGCAAC	660	
Qy	661	TCGGTTCAGGCCCTGGCAATCCCAACCAAAATCATCCTGAATCTGATCTTTTTATACAA	720	
Db	661	TCGGTTCAGGCCCTGGCAATCCCAACCAAAATCATCCTGAATCTGATCTTTTTATACAA	720	

Qy	661	TCGGT	CAGCCCC	TGGCAAT	CCCA	CCAAT	CAT	CCT	GAAT	CT	GAT	CT	TTTT	TAT	CA	CAAT	720
Db	661	TCGGT	CAGCCCC	CTGGCAAT	CCCA	CAAT	CAT	CCT	GAAT	CT	GAT	CT	TTTT	TAT	CA	CAAT	720

QY 721 ATACGAAAGCCAGCTTGAA 740
Db 721 ATACGAAAGCCAGCTTGAA 740

RESULT 5
US-09-841-879B-19/c
; Sequence 19, Application US/09841879B
; Patent No. US20020142979A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-330001
; CURRENT APPLICATION NUMBER: US/09/841,879B
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: US 09/728,721
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-841-879B-19

Query Match 100.0%; Score 740; DB 10; Length 740;
Best Local Similarity 100.0%; Pred. No. 1.1e-196;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGTCCGGCTGACGGGGTGAGCGCGGCGAGCGCGCGGGATCCTGGAGCATGGGCG 60
Db 740 CGCGTCCGGCTGACGGGGTGAGCGCGGCGAGCGCGCGGGATCCTGGAGCATGGGCG 681

QY 61 GCGCGCGCGACGCATCTCGATGCGCTGGAGAACCTGACCGCGGAGGAGCTCAAGAAGT 120
Db 680 GCGCGCGCGACGCATCTCGATGCGCTGGAGAACCTGACCGCGGAGGAGCTCAAGAAGT 621

QY 121 TCAGCTGAAGTGTGTGCTGCGTCCGCTGCGGAGGGCTACGGCGGCATCCCGCGGGCG 180
Db 620 TCAAGCTGAAGTGTGTGCTGCGTCCGCTGCGGAGGGCTACGGCGGCATCCCGCGGGCG 561

QY 181 CGTGTGTCCTAGGAGCGCTTGGACCTCACCGACAAGCTGTGACGTTCTACCTGGAGA 240
Db 560 CGTGTGTCCTAGGAGCGCTTGGACCTCACCGACAAGCTGTGACGTTCTACCTGGAGA 501

QY 241 CCTACGGCGCGAGCTCACCGCTAACGCTGTGCGGACATGGGCTTGCAGGAGATGGCGG 300
Db 500 CCTACGGCGCGAGCTCACCGCTAACGCTGTGCGGACATGGGCTTGCAGGAGATGGCGG 441

QY 301 GCGAGCTGACGGGGCGACGACAGGCGCTTGGAGCGCGCGAGCTGGGATCCAGGCGCC 360
Db 440 GCGAGCTGACGGGGCGACGACAGGCGCTTGGAGCGCGCGAGCTGGGATCCAGGCGCC 381

QY 361 CTCCTCAGTCGCGAGCCAGCGCGCTGACCTTTATAGACCAGCAGCGGGCTGCGCTTA 420
Db 380 CTCCTCAGTCGCGAGCCAGCGCGCTGACCTTTATAGACCAGCAGCGGGCTGCGCTTA 321

QY 421 TCGGAGGGTCAAAACGTTGAGTGGCTGCTGGATGCTTGTACGGGAAGGCTCTGACGCG 480
Db 320 TCGGAGGGTCAAAACGTTGAGTGGCTGCTGGATGCTTGTACGGGAAGGCTCTGACGCG 261

QY 481 ATGAGCAGTACAGGAGTGGGGCGGCGACCAACCAAGCAAGATGCGGAAGCTCT 540
Db 260 ATGAGCAGTACAGGAGTGGGGCGGCGACCAACCAAGCAAGATGCGGAAGCTCT 201

QY 541 TCAGTTTACACAGCTGGAAGTGGAGCTGCTTCTTCCAGGGCCCTAAGG 600
Db 200 TCAGTTTACACAGCTGGAAGTGGAGCTGCTTCTTCCAGGGCCCTAAGG 141

QY 601 AGTCCAGTCTTACCTGTGGAGGACCTGGAGCGAGCTGAGGCTCTTCCCGCAACAC 660
Db 601 AGTCCAGTCTTACCTGTGGAGGACCTGGAGCGAGCTGAGGCTCTTCCCGCAACAC 660

Db 140 AGTCCAGTCTTACCTGTGGAGGACCTGGAGCGAGGCTCTTCCCGCAACAC 81
QY 661 TCCGCTCAGCCCTGGCAATCCCAAAATCATCTGAATCTGATCTTTTATACACAAT 720
Db 80 TCCGCTCAGCCCTGGCAATCCCAAAATCATCTGAATCTGATCTTTTATACACAAT 21

QY 721 ATACGAAAGCCAGCTTGAA 740
Db 20 ATACGAAAGCCAGCTTGAA 1

RESULT 6
US-10-295-981-48
; Sequence 48, Application US/10295981
; Publication No. US20030120055A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/10/295,981
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US/09/340,620
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)...(638)
US-10-295-981-48

Query Match 100.0%; Score 740; DB 15; Length 740;
Best Local Similarity 100.0%; Pred. No. 1.1e-196;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGTCCGGCTGACGGGGTGAGCGCGGCGAGCGCGGGATCCTGGAGCATGGGCG 60
Db 1 CGCGTCCGGCTGACGGGGTGAGCGCGGCGAGCGCGGGATCCTGGAGCATGGGCG 60

QY 61 GCGCGCGGAGCCATCTGGATGCGCTGCGGAGGCTACGGGCGCATCCCGCGGGCG 120
Db 61 GCGCGCGGAGCCATCTGGATGCGCTGCGGAGGCTACGGGCGCATCCCGCGGGCG 120

QY 121 TCAAGCTGAAGTGTGCTGCGTCCGCTGCGGAGGCTACGGGCGCATCCCGCGGGCG 180
Db 121 TCAAGCTGAAGTGTGCTGCGTCCGCTGCGGAGGCTACGGGCGCATCCCGCGGGCG 180

QY 181 CGTGTCTTCATGACCGCTTGGACCTCACCGACAAGCTGGTTCAGCTTCTACCTGGAGA 240
Db 181 CGTGTCTTCATGACCGCTTGGACCTCACCGACAAGCTGGTTCAGCTTCTACCTGGAGA 240

QY 241 CCTACGGCGCGAGCTCACCGCTAACGCTGCGGACATGGGCTGCGAGAGATGGCGG 300
Db 241 CCTACGGCGCGAGCTCACCGCTAACGCTGCGGACATGGGCTGCGAGAGATGGCGG 300

QY 301 GCGAGCTCAGCGCGCCACGACACAGGGCTCTTGGAGCGCGCGAGCTGGGATCCAGGCC 360
Db 301 GCGAGCTCAGCGCGCCACGACACAGGGCTCTTGGAGCGCGCGAGCTGGGATCCAGGCC 360

QY 361 CTCCTCAGTCGCGAGCCCAAGCCAGGCTGCACTTTATAGACAGCAGCCGGGCTGCGCTTA 420
Db 361 CTCCTCAGTCGCGAGCCCAAGCCAGGCTGCACTTTATAGACAGCAGCCGGGCTGCGCTTA 420

QY 421 TCCGAGGCTCAAAACGTTGAGTGGCTCTGTGATGCTCTGTACGGAAAGGTCCTTGACGG 480
Db 421 TCCGAGGCTCAAAACGTTGAGTGGCTCTGTGATGCTCTGTACGGAAAGGTCCTTGACGG 480
QY 481 ATGACGAGTACAGGAGTGGCGGCGGAGCCACCAACCCCAAGCAAGATGCGGAAGCTCT 540
Db 481 ATGACGAGTACAGGAGTGGCGGCGGAGCCACCAACCCCAAGCAAGATGCGGAAGCTCT 540
QY 541 TCAGTTTACACAGCCTCGAACTGGACCTGCAAGGACTTGTCTCTCCAGGCGCTAAAGG 600
Db 541 TCAGTTTACACAGCCTCGAACTGGACCTGCAAGGACTTGTCTCTCCAGGCGCTAAAGG 600
QY 601 AGTCCAGTCTTACCTGTGTGAGGAGCTTGGAGCGGAGCTGAGGCTCTTCCAGCAACAC 660
Db 601 AGTCCAGTCTTACCTGTGTGAGGAGCTTGGAGCGGAGCTGAGGCTCTTCCAGCAACAC 660
QY 661 TCCGTCAGCCCTGCAATCCCAACCAATCATCTGAACTCTGATCTTTTATACAACT 720
Db 661 TCCGTCAGCCCTGCAATCCCAACCAATCATCTGAACTCTGATCTTTTATACAACT 720
QY 721 ATACGAAAAGCAGCTTGAA 740
Db 721 ATACGAAAAGCAGCTTGAA 740
RESULT 7
US-10-240-145-12
; Sequence 12, Application US/10240145
; Publication No. US20030235883A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-048
; CURRENT APPLICATION NUMBER: US/10/240,145
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/668,680
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/695,618
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 09/728,711
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: NOT YET ASSIGNED
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: Custom
; SEQ ID NO 12
; LENGTH: 745
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (48)..(632)
US-10-240-145-12
Query Match 99.1%; Score 733; DB 12; Length 745;
Best Local Similarity 100.0%; Pred. No. 9.7e-195;
Matches 733; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 8 GCGTCAGCGGGGTGACGCGCGCAGCGCGGGATCTGGAGCCATGGGGCGCGCG 67
Db 2 GCGTCAGCGGGGTGACGCGCGCAGCGCGGGATCTGGAGCCATGGGGCGCGCG 61
QY 68 CGACGCCATCTCGATCGCTGGAGAACCTGACCGCGGAGGCTCAAGAAAGTTCAAGCT 127
Db 62 CGACGCCATCTCGATCGCTGGAGAACCTGACCGCGGAGGCTCAAGAAAGTTCAAGCT 121
QY 128 GAAGCTGCTGTGCGGTCCGCTGCGAGGAGCTACGCGCGCATCCCGCGGGCGCGCTGCT 187

Db 122 GAAGCTGCTGTGCGGTCCGCTGCGCGAGGGCTACGCGCGCATCCCGGGGGCGCGCTGCT 181
QY 188 GTCCATGGACGCTTGGACCTCACCGCAAGCTGTGTAGCTTTCTACCTGGAGACCTACGG 247
Db 182 GTCCATGGACGCTTGGACCTCACCGCAAGCTGTGTAGCTTTCTACCTGGAGACCTACGG 241
QY 248 CGCCGAGCTCACCGCTAACGCTGCTGCGGACATGGGCTTGCAGGAGATGGCGGGCAGCT 307
Db 242 CGCCGAGCTCACCGCTAACGCTGCTGCGGACATGGGCTTGCAGGAGATGGCGGGCAGCT 301
QY 308 CGAGCGCGCCACCGCACCGAGGCTCTGGAGCCGCGCAGCTGGGATCCAGGCCCTCTCTCA 367
Db 302 CGAGCGCGCCACCGCACCGAGGCTCTGGAGCCGCGCAGCTGGGATCCAGGCCCTCTCTCA 361
QY 368 GTCCGAGCCCAAGCCAGGCGCTGACCTTTATAGACCAAGCAGCGGGCTTGTATCGCGAG 427
Db 362 GTCCGAGCCCAAGCCAGGCGCTGACCTTTATAGACCAAGCAGCGGGCTTGTATCGCGAG 421
QY 428 GGTCAAAACGTTGAGTGGCTGCTGATGCTGTACGGAAAGGTCCTGACGGATGAGCA 487
Db 422 GGTCAAAACGTTGAGTGGCTGCTGATGCTGTACGGAAAGGTCCTGACGGATGAGCA 481
QY 488 GTACCAGCAGTGGCGGCGGAGCCACCAACCAAGCAAGATGCGGAAGCTCTTCAGTTT 547
Db 482 GTACCAGCAGTGGCGGCGGAGCCACCAACCAAGCAAGATGCGGAAGCTCTTCAGTTT 541
QY 548 CACACAGCCTGGAACTGGACCTGCAAGGACTTGTCTCTCCAGGGCCTTAAGGAGTCCCA 607
Db 542 CACACAGCCTGGAACTGGACCTGCAAGGACTTGTCTCTCCAGGGCCTTAAGGAGTCCCA 601
QY 608 GTCTTACTGTGTGAGGACCTTGGAGCGGAGCTGAGGCTCTTCCAGCAACACTCGGGTC 667
Db 602 GTCTTACTGTGTGAGGACCTTGGAGCGGAGCTGAGGCTCTTCCAGCAACACTCGGGTC 661
QY 668 AGCCCTGGCAATCCCAACCAATCATCTGATCTGATCTTTTATACAAATATACGAA 727
Db 662 AGCCCTGGCAATCCCAACCAATCATCTGATCTGATCTTTTATACAAATATACGAA 721
QY 728 AAGCCAGCTTGAA 740
Db 722 AAGCCAGCTTGAA 734
RESULT 8
US-10-131-410-21
; Sequence 21, Application US/10131410
; Publication No. US20030235915A1
; GENERAL INFORMATION:
; APPLICANT: SPECHT, THOMAS
; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILARSKY, CHRISTIAN
; APPLICANT: DAHL, EDGAR
; APPLICANT: ROSENTHAL, ANDRE
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM TISSUE OF BREAST
; TITLE OF INVENTION: TUMORS
; FILE REFERENCE: SCH-1763
; CURRENT APPLICATION NUMBER: US/10/131,410
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 09/646,673
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: PCT/DE99/00908
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 202
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 779
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-131-410-21
Query Match 99.1%; Score 733; DB 12; Length 779;
Best Local Similarity 100.0%; Pred. No. 9.8e-195;

; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 278
; LENGTH: 811
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-301-278

Query Match 94.3%; Score 697.6; DB 9; Length 811;
Best Local Similarity 99.5%; Pred. No. 7.6e-185;
Matches 730; Conservative 1; Mismatches 0; Indels 3; Gaps 3;

QY 7 CGGCTGCAGCGGGTGCAGCGGCGGAGCCTCGAGCCATGGGGCGGCGC 66
DB 68 CGGCTGCAGCGGGTGCAGCGGCGGAGCCTCGAGCCATGGGGCGGCGC 127
QY 67 GCGAGCGCATCTGGATGGCTGGAGAACCTGACCGCGGAGGAGCTCAAGAAAGTTCAAGC 126
DB 128 GCGAGCGCATCTGGATGGCTGGAGAACCTGACCGCGGAGGAGCTCAAGAAAGTTCAAGC 187
QY 127 TGAAGCTGTGTGGTGGCGGCTGCGGAGGGCTACGCGGCGCATCCCGGGGGCGGCTGC 186
DB 188 TGAAGCTGTGTGGTGGCGGCTGCGGAGGGCTACGCGGCGCATCCCGGGGGCGGCTGC 247
QY 187 TGTCCATGACGCTTGGACCTTACCGACAAAGTGTGAGCTTCTACCTGGAGACCTAGC 246
DB 248 TGTCCATGACGCTTGGACCTTACCGACAAAGTGTGAGCTTCTACCTGGAGACCTAGC 307
QY 247 GCGCGGAGCTCACCGCTAACGCTGCTGCGGACATGGGCGCTGAGGAGATGGCGGCGCAGC 306
DB 308 GCGCGGAGCTCACCGCTAACGCTGCTGCGGACATGGGCGCTGAGGAGATGGCGGCGCAGC 367
QY 307 TGCAGCGCGCCACGACCGAGGCTGTGAGCGCGCGCAGCTGGGATCCAGGCGCCCTCCTC 366
DB 368 TGCAGCGCGCCACGACCGAGGCTGTGAGCGCGCGCAGCTGGGATCCAGGCGCCCTCCTC 427
QY 367 AGTGGCAGCGCAGCGGCGCTGCACTTTATAGACCGACCGGGCTGGCTTATCGGA 426
DB 428 AGTGGCAGCGCAGCGGCGCTGCACTTTATAGACCGACCGGGCTGGCTTATCGGA 487
QY 427 GGTCAACAAAGTTGAGTGGCTGCTGATGCTGTGACGGAGGCTTCAAGGATGAGC 486
DB 488 GGTCAACAAAGTTGAGTGGCTGCTGATGCTGTGACGGAGGCTTCAAGGATGAGC 547
QY 487 AGTACCGAGCAGTGGCGGCGGAGCCACCAACCAAGCAAGATGCGGAAGCTCTTCAGTT 546
DB 548 AGTACCGAGCAGTGC-GGCGGAGCCACCAACCAAGCAAGATGCGGAAGCTCTTCAGTT 606
QY 547 TCACACGAGCTGGAACCTGGAACCTGGAACCTTGTCTCTCCAGGCCCTAAGGGAGTCCC 606
DB 607 TCACACGAGCTGGAACCTGGAACCTTGTCTCTCCAGGCCCTAAGGGAGTCCC 666
QY 607 AGTCTTACCTGTGGAGGACCTGGAGCGGAGCTGAGGCTCCTTCCAGCAACACTCCCGT 666
DB 667 AGTCTTACCTGTGGAGGACCTGGAGC-GAGCTGAGGCTCCTTCCAGCAACACTCCCGT 725
QY 667 CAGCCCTGGCAATCCCAACAAATCATCTGAATCTGATCTTTTATACAAATATACGA 726
DB 726 CA-SCCTGGCAATCCCAACAAATCATCTGAATCTGATCTTTTATACAAATATACGA 784
QY 727 AAAGCCAGCTTGAA 740
DB 785 AAAGCCAGCTTGAA 798

RESULT 11
US-09-728-721-50
; Sequence 50, Application US/09728721
; Patent No. US20020061845A1

; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/728,721
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 585
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-728-721-50

Query Match 79.1%; Score 585; DB 9; Length 585;
Best Local Similarity 100.0%; Pred. No. 2e-153;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 ATGGGCGCGCGCGCGAGCCATCTCGATCGCTGGAGAACCTGACCGCGGAGGATC 113
DB 1 ATGGGCGCGCGCGCGAGCCATCTCGATCGCTGGAGAACCTGACCGCGGAGGATC 60
QY 114 AAGAAAGTTCAAGCTGAAGCTGTGTGCTGCGCTGCGGAGGCTACGGGCGCATCCCG 173
DB 61 AAGAAAGTTCAAGCTGAAGCTGTGTGCTGCGCTGCGGAGGCTACGGGCGCATCCCG 120
QY 174 CGGCGCGCGCTGCTGCTCATGAGCGCTTGGACCTGACCGACCAAGCTGTCAGCTTCTAC 233
DB 121 CGGCGCGCGCTGCTGCTCATGAGCGCTTGGACCTGACCGACCAAGCTGTCAGCTTCTAC 180
QY 234 CTGAGAGACTACGGCGCGAGCTCACCGCTAACGCTGCTGCGGACATGGGCGCTGAGGAG 293
DB 181 CTGAGAGACTACGGCGCGAGCTCACCGCTAACGCTGCTGCGGACATGGGCGCTGAGGAG 240
QY 294 ATGGCGCGGAGCTGAGCGCGGCGCACGACCAAGGCTCTGAGAGCGCGCGCAGCTGGGATC 353
DB 241 ATGGCGCGGAGCTGAGCGCGGCGCACGACCAAGGCTCTGAGAGCGCGCGCAGCTGGGATC 300
QY 354 CAGGCGCGCTCTCAGTGGGAGCGGAGCGGAGCGCTGCACTTTATAGACCGACCGGCT 413
DB 301 CAGGCGCGCTCTCAGTGGGAGCGGAGCGGAGCGGAGCGCTGCACTTTATAGACCGACCGGCT 360
QY 414 GCGCTTATCGGAGGCTCACAAACGTTGAGTGGCTGCTGGATGCTCTGTACGGGAGGTC 473
DB 361 GCGCTTATCGGAGGCTCACAAACGTTGAGTGGCTGCTGGATGCTCTGTACGGGAGGTC 420
QY 474 CTGACGGATGAGCAGTACAGGAGCTGCGGGCGGAGCGGAGCGGAGCGGAGGATCGG 533
DB 421 CTGACGGATGAGCAGTACAGGAGCTGCGGGCGGAGCGGAGCGGAGCGGAGGATCGG 480
QY 534 AAGCTCTTCAGTTTACACCGAGCTTGGAACTGGAGCTGCAAGGACTTGTCTCTCCAGGCT 593
DB 481 AAGCTCTTCAGTTTACACCGAGCTTGGAACTGGAGCTGCAAGGACTTGTCTCTCCAGGCT 540
QY 594 CTAAGGGAGTCCAGTCTTACCTGCTGGAGGAGCTTGGAGCGGAGC 638
DB 541 CTAAGGGAGTCCAGTCTTACCTGCTGGAGGAGCTTGGAGCGGAGC 585

RESULT 12
US-09-996-617-9
; Sequence 9, Application US/09996617
; Patent No. US20020128198A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John

; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; FILE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-340001
; CURRENT APPLICATION NUMBER: US/09/996,617
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: 09/931,071
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/428,252
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 585
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-996-617-9

Query Match 79.1%; Score 585; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 2e-153;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 ATGGGCGCGCGCGAAGCCATCTCTGGATGGCTGGAGAACCTGACCGCGAGGAGCTC 113
DB 1 ATGGGCGCGCGCGAAGCCATCTCTGGATGGCTGGAGAACCTGACCGCGAGGAGCTC 60
QY 114 AAGAAGTTCAAGCTGAAGCTGCTGCTGGTGGCTGGAGGAGCTACGGGCGCATCCCG 173
DB 61 AAGAAGTTCAAGCTGAAGCTGCTGCTGGTGGCTGGAGGAGCTACGGGCGCATCCCG 120
QY 174 CGGGCGCGCTGCTGCTCAATGAGAGCGCTTGGACCTTGGACCAAGCTGGTCTTCTAC 233
DB 121 CGGGCGCGCTGCTGCTCAATGAGAGCGCTTGGACCTTGGACCAAGCTGGTCTTCTAC 180
QY 234 CTGGAGACCTACGGCGCGAGCTCACCGCTAAACGTTGCTGCGGACATGGGCTGAGGAG 293
DB 181 CTGGAGACCTACGGCGCGAGCTCACCGCTAAACGTTGCTGCGGACATGGGCTGAGGAG 240
QY 294 ATGGCGGGGAGCTGACGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGATC 353
DB 241 ATGGCGGGGAGCTGACGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGATC 300
QY 354 CAGGCGCCCTCTCTCAGTCGGGAGCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 413
DB 301 CAGGCGCCCTCTCTCAGTCGGGAGCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCT 360
QY 414 GCGCTTATCGGAGGGTCAAAAAGTTGAGTGGCTGCTGGATGCTCTTACGGGAGGTC 473
DB 361 GCGCTTATCGGAGGGTCAAAAAGTTGAGTGGCTGCTGGATGCTCTTACGGGAGGTC 420
QY 474 CTGACGGATGAGCAGTACAGGCGGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGAG 533
DB 421 CTGACGGATGAGCAGTACAGGCGGAGTGGGCGGCGGCGGCGGCGGCGGCGGAG 480
QY 534 AAGCTCTTCAAGTTTCAACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 593
DB 481 AAGCTCTTCAAGTTTCAACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
QY 594 CTAAGGGAGTCCAGTCTTCACTGGTGGAGGAGCTTGGAGCGGAGC 638
DB 541 CTAAGGGAGTCCAGTCTTCACTGGTGGAGGAGCTTGGAGCGGAGC 585

RESULT 13
US-09-841-879B-6
; Sequence 6, Application US/09841879B
; Patent No. US20020142979A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-330001
; CURRENT APPLICATION NUMBER: US/09/841,879B

; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: US 09/728,721
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 585
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-841-879B-6

Query Match 79.1%; Score 585; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 2e-153;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 ATGGGCGCGCGCGAGCGCCATCTCTGGATGGCTGGAGAACCTGACCGCGAGGAGCTC 113
DB 1 ATGGGCGCGCGCGAGCGCCATCTCTGGATGGCTGGAGAACCTGACCGCGAGGAGCTC 60
QY 114 AAGAAGTTCAAGCTGAAGCTGCTGCTGGTGGCTGGAGGAGCTACGGGCGCATCCCG 173
DB 61 AAGAAGTTCAAGCTGAAGCTGCTGCTGGTGGCTGGAGGAGCTACGGGCGCATCCCG 120
QY 174 CGGGCGCGCTGCTGCTCAATGAGAGCGCTTGGACCTTGGACCAAGCTGGTCTTCTAC 233
DB 121 CGGGCGCGCTGCTGCTCAATGAGAGCGCTTGGACCTTGGACCAAGCTGGTCTTCTAC 180
QY 234 CTGGAGACCTACGGCGCGAGCTCACCGCTAAACGTTGCTGCGGACATGGGCTGAGGAG 293
DB 181 CTGGAGACCTACGGCGCGAGCTCACCGCTAAACGTTGCTGCGGACATGGGCTGAGGAG 240
QY 294 ATGGCGGGGAGCTGACGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGATC 353
DB 241 ATGGCGGGGAGCTGACGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGATC 300
QY 354 CAGGCGCCCTCTCTCAGTCGGGAGCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 413
DB 301 CAGGCGCCCTCTCTCAGTCGGGAGCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCT 360
QY 414 GCGCTTATCGGAGGGTCAAAAAGTTGAGTGGCTGCTGGATGCTCTTACGGGAGGTC 473
DB 361 GCGCTTATCGGAGGGTCAAAAAGTTGAGTGGCTGCTGGATGCTCTTACGGGAGGTC 420
QY 474 CTGACGGATGAGCAGTACAGGCGGAGTGGGCGGCGGCGGCGGCGGCGGCGGAG 533
DB 421 CTGACGGATGAGCAGTACAGGCGGAGTGGGCGGCGGCGGCGGCGGCGGCGGAG 480
QY 534 AAGCTCTTCAAGTTTCAACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 593
DB 481 AAGCTCTTCAAGTTTCAACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
QY 594 CTAAGGGAGTCCAGTCTTCACTGGTGGAGGAGCTTGGAGCGGAGC 638
DB 541 CTAAGGGAGTCCAGTCTTCACTGGTGGAGGAGCTTGGAGCGGAGC 585

RESULT 14
US-10-295-981-50
; Sequence 50, Application US/10295981
; Publication No. US20030120055A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/10/295,981
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US/09/340,620
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; CURRENT APPLICATION NUMBER: US 09/207,359

; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-981-50

Query Match 79.1%; Score 585; DB 15; Length 585;
Best Local Similarity 100.0%; Pred. No. 2e-153;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 ATGGGCGCGCGCGACGCATCTTGGATGGCTGAGAACCTGACCGCCGAGGAGTTC 113
Db 1 ATGGGCGCGCGCGACGCATCTTGGATGGCTGAGAACCTGACCGCCGAGGAGTTC 60
QY 114 AAGAAGTTCAAGCTGAAGCTGCTGTCGGTGCCTGCGGAGGGCTACCGGCGCATCCCG 173
Db 61 AAGAAGTTCAAGCTGAAGCTGCTGTCGGTGCCTGCGGAGGGCTACCGGCGCATCCCG 120
QY 174 CGGGGCGCGCTGCTGTCATGGAAGCTTGGACCTTCAACGACAAAGCTGCTCAGCTTCTAC 233
Db 121 CGGGGCGCGCTGCTGTCATGGAAGCTTGGACCTTCAACGACAAAGCTGCTCAGCTTCTAC 180
QY 234 CTGGAGACCTACCGGCGGAGCTCAACGCTAAGCTGCTGCGGACATGGGCTCGAGGAG 293
Db 181 CTGGAGACCTACCGGCGGAGCTCAACGCTAAGCTGCTGCGGACATGGGCTCGAGGAG 240
QY 294 ATGGCGCGGAGCTGACGAGCGGCGACGACAGGGGCTCTGGAGCGCGCAGCTGGGATC 353
Db 241 ATGGCGCGGAGCTGACGAGCGGCGACGACAGGGGCTCTGGAGCGCGCAGCTGGGATC 300
QY 354 CAGGCGCCCTCCTCAGTGGGAGCAAGCGGCTGCTGATAGACAGCAACCGGGCT 413
Db 301 CAGGCGCCCTCCTCAGTGGGAGCAAGCGGCTGCTGATAGACAGCAACCGGGCT 360
QY 414 GCGCTTATCGGAGGCTCAAAAGTGTAGTGGCTGCTGATGCTCTGTACGGGAAGTTC 473
Db 361 GCGCTTATCGGAGGCTCAAAAGTGTAGTGGCTGCTGATGCTCTGTACGGGAAGTTC 420
QY 474 CTGACGGATGAGCAGTACAGGAGTGGCGGCGAGCGCCAGCCACCAAGCAAGATGCGG 533
Db 421 CTGACGGATGAGCAGTACAGGAGTGGCGGCGAGCGCCAGCCACCAAGCAAGATGCGG 480
QY 534 AAGCTCTTCAGTTTCAACAGCCTGGAACTGGACCTGCAAGGACTTGTCTCCAGGGCC 593
Db 481 AAGCTCTTCAGTTTCAACAGCCTGGAACTGGACCTGCAAGGACTTGTCTCCAGGGCC 540
QY 594 CTAAGGGAGTCCAGTCCCTACCTGTTGGAGGAGCTGGAGCGGAGCTTGAAG 638
Db 541 CTAAGGGAGTCCAGTCCCTACCTGTTGGAGGAGCTGGAGCGGAGCTTGAAG 585

RESULT 15
US-10-106-698-1145
; Sequence 1145, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280

; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 1145
; LENGTH: 639
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-106-698-1145

Query Match 71.8%; Score 531; DB 15; Length 639;
Best Local Similarity 94.9%; Pred. No. 2.4e-138;
Matches 595; Conservative 0; Mismatches 20; Indels 12; Gaps 4;

QY 124 AGCTGAAGCTGCTGCTGCGGCTGCGGAGGGCTACGGGCGCATCCCGG-----G 176
Db 2 AGCTGAAGCTGCTGCTGCGGCTGCGGAGGGCTACGGGCGCGCGACGCCATCTCTG 61
QY 177 GCGCGCGCTGCTTCATGAGCGCCTTGACCTCACCGA---CAAGCTGGTCAAGCTTCTAC 233
Db 62 GATGCGCTGGAGAACCTGACCGCCGAGGAGCTCAAGAAAGTTCAAGCTGGTCAAGCTTCTAC 121
QY 234 CTGGAGACCTACCGGCGCGAGCTCAACGCTAAGCTGCTGCGGACATGGGCTCGAGGAG 293
Db 122 CTGGAGACCTACCGGCGCGAGCTCAACGCTAAGCTGCTGCGGACATGGGCTCGAGGAG 181
QY 294 ATGGCGCGGAGCTGACGAGCGGCGACGACCAAGGGCTCTGGAGCGCGCAGCTGGGATC 353
Db 182 ATGGCGCGGAGCTGACGAGCGGCGACGACCAAGGGCTCTGGAGCGCGCAGCTGGGATC 241
QY 354 CAGGCGCCCTCCTCAGTGGGAGCAAGCGGCTGCTGATAGACAGCAACCGGGCT 413
Db 242 CAGGCGCCCTCCTCAGTGGGAGCAAGCGGCTGCTGATAGACAGCAACCGGGCT 301
QY 414 GCGCTTATCGGAGGCTCACAAAGTGTAGTGGCTGCTGATGCTCTGTACGGGAAGTTC 473
Db 302 GCGCTTATCGGAGGCTCACAAAGTGTAGTGGCTGCTGATGCTCTGTACGGGAAGTTC 361
QY 474 CTGACGGATGAGCAGTACAGGAGTGGCGGCGAGCGCCACCAAGCAAGATGCGG 533
Db 362 CTGACGGATGAGCAGTACAGGAGTGGCGGCGAGCGCCACCAAGCAAGATGCGG 420
QY 534 AAGCTCTTCAGTTTCAACAGCCTGGAACTGGACCTGCAAGGACTTGTCTCCAGGGCC 593
Db 421 AAGCTCTTCAGTTTCAACAGCCTGGAACTGGACCTGCAAGGACTTGTCTCCAGGGCC 480
QY 594 CTAAGGGAGTCCAGTCCCTACCTGTTGGAGGAGCTGGAGCGGAGCTTGAAG 653
Db 481 CTAAGGGAGTCCAGTCCCTACCTGTTGGAGGAGCTGGAGCGGAGCTTGAAG 540
QY 654 GCAACACCTCGGCTCAGCCCTGGCAATCCCAACCAATCATCTGAACTCTGATCTTTTAT 713
Db 541 GCAACACCTCGGCTCAG-CCCTGGCAATCCCAACCAATCATCTGAACTCTGATCTTTTAT 599
QY 714 ACACAATATACAAAGAGCAGCTTGAA 740
Db 600 ACACAATATACAAAGAGCAGCTTGAA 626

Search completed: January 29, 2004, 13:44:37
Job time : 236.019 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 03:29:26 ; Search time 9.48584 Seconds
(without alignments) 3262.936 Million cell updates/sec

Title: US-09-996-617-8

Perfect score:

Sequence: 1 MGRARDAILDALENLTAEEL.....LLQLALRESQSYLVEDLERS 195

Scoring table: BLOSUM62

```

scoring table: BEO30M02
                Gapop 10.0 , Gapext 0.5

```

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Minimum	DB seq	length:
Maximum	DB seq	length: 2000000000

post-processing: Minimum Match 0%

Fast-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 19Jun03:*

1:	/SIDSI1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2:	/SIDSI1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3:	/SIDSI1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4:	/SIDSI1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5:	/SIDSI1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6:	/SIDSI1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7:	/SIDSI1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8:	/SIDSI1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9:	/SIDSI1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10:	/SIDSI1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11:	/SIDSI1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12:	/SIDSI1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13:	/SIDSI1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14:	/SIDSI1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15:	/SIDSI1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16:	/SIDSI1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17:	/SIDSI1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18:	/SIDSI1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19:	/SIDSI1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20:	/SIDSI1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21:	/SIDSI1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22:	/SIDSI1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23:	/SIDSI1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*
24:	/SIDSI1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query			DB	ID	Description
		Match	Length	Length			
1	990	100.0	195	20	AAU48553	Human breast tumou	
2	990	100.0	195	22	AAU68525	Human novel cytoki	
3	990	100.0	195	22	AAE00588	Human target of me	
4	990	100.0	195	22	AAE20085	Human CARD-5 prote	
5	990	100.0	195	23	AAU93935	Human caspase recr	
6	990	100.0	195	23	AAO17854	Pyrin domain conta	
7	990	100.0	195	24	ABG71635	Human caspase recr	
8	990	100.0	205	22	AAG74647	Human colon cancer	
9	873.5	88.2	176	22	AAE00594	Alternatively spli	

10	759	76.7	190	21	AAB343675	Human cancer assoc
11	684	69.1	193	22	AAB00592	Mouse target of me
12	684	69.1	193	22	AAB00592	Mouse CARD-5 prote
13	684	69.1	193	23	AAB00592	Mouse caspase recr
14	576	58.2	171	22	AAB00593	Rat target of meth
15	564.5	57.0	158	22	AAG74648	Human colon cancer
16	545	55.1	110	22	AAB00595	Alternatively spli
17	509	51.4	136	21	AAB53861	Human colon cancer
18	448	45.3	91	22	AAB00599	Human target of me
19	439	44.3	84	22	AAB00591	Human target of me
20	350	30.8	76	23	AAO21934	PAAD domain-contai
21	267	27.0	89	22	AAB36608	Human FLEXHT-30 pr
22	267	27.0	89	23	AAO17852	Pyrin domain conta
23	253	25.6	1397	22	AAI72670	Human NB-ARC and C
24	250.5	25.3	1429	22	AAB62571	Human CARD-7 polyp
25	250.5	25.3	1429	23	ABG37969	Human leucine rich
26	250.5	25.3	1429	23	ABG78472	Leucine-rich repea
27	250.5	25.3	1429	23	ABB77916	Human leucine-rich
28	250.5	25.3	1429	24	ABG71631	Human caspase recr
29	250.5	25.3	1429	24	ABG71633	Human caspase recr
30	250	25.3	1442	22	AAI72671	Human NB-ARC and C
31	250	25.3	1473	22	AAI72669	Human NB-ARC and C
32	250	25.3	1473	22	AAI72711	Human NAC beta iso
33	249.5	25.2	1429	23	ABG78455	Human caspase recr
34	249.5	25.2	1429	23	AAO17855	Human domain conta
35	248.5	25.1	1473	22	AAB06758	Pyrin domain conta
36	240	24.2	442	21	AAB24513	Human secreted pro
37	182	18.4	65	23	AAO21935	PAAD domain-contai
38	153	15.5	65	23	AAO24519	Human secreted pro
39	147	14.8	77	22	AAB07516	Human PYRIN-1 doma
40	137	13.8	77	22	AAB07515	Human PYRIN-1 doma
41	125.5	12.7	891	22	AAB04546	Human G-protein coo
42	125.5	12.7	1851	22	AAU00023	Human activated T-
43	125.5	12.7	1851	23	AAO17860	Pyrin domain conta
44	122.5	12.4	781	20	AAI09001	FMF associated pro
45	122.5	12.4	781	23	AAO17853	Pyrin domain conta

ALIGNMENTS

RESULT 1	
AAV48553	
ID	AAV48553 standard; Protein; 195 AA.
XX	
AC	AAV48553;
XX	
DT	08-DEC-1999 (first entry)
XX	
DE	Human breast tumour-associated protein 14.
XX	
KW	Expressed sequence tag; EST; human; breast; cancer; gene therapy; treatment; tumour; cytostatic; medicament.
XX	
XX	
OS	Homo sapiens.
XX	
PN	DE19813839-A1.
XX	
PD	23-SEP-1999.
XX	
PF	20-MAR-1998; 98DE-1013839.
XX	
PR	20-MAR-1998; 98DE-1013839.
XX	
XX	
PA	(META-) METAGEN GES GENOMFORSCHUNG MBH.
XX	
PI	Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenka
XX	
DR	WPI; 1999-528981/45.
DR	N-PSDB; AA23631.
XX	
PT	Human nucleic acid sequences and protein products from tumor breas
XX	
PT	tissue, useful for breast cancer therapy -

Human nucleic acid sequences and protein products from tumor breast tissue, useful for breast cancer therapy -

AAE00588;
02-JUL-2001 (first entry)
Human target of methylation-induced silencing-1 (TMS1) protein.
Human; target of methylation-induced silencing-1; TMS1; cytostatic;
anti-proliferative; apoptosis inducer; gene therapy; CpG island;
caspase-recruiting domain; CARD; cancer; breast.
Homo sapiens.
WO200129235-A2.
26-APR-2001.
18-OCT-2000; 2000WO-US28747.
18-OCT-1999; 99US-0159975.
(UYEM-) UNIV EMORY.
Vertino PM;
WPI; 2001-290922/30.
N-PSDB; AAD03889, AAD03890.
Novel gene TMS1, transcriptionally silenced due to increased
methylation useful for identifying subject at risk of developing tumor
characterized by abnormal methylation, for treating cancer by inducing
apoptosis -
Claim 85; Page 114; 124pp; English.
The invention relates to identification of target of methylation-induced
silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to
abnormal methylation of a CpG island in its 5' regulatory region. TMS1
consists of a carboxy terminal caspase-recruiting domain (CARD) and
plays a role in induction of apoptosis. TMS1 gene and protein are useful
as tools for diagnosing and treating a subject at risk of developing
cancer (e.g. breast cancer) characterized by abnormal CpG methylation or
abnormally low levels of TMS1 expression products. Unique fragments of
TMS1 molecule are used as probes. TMS1 gene is useful in gene therapy.
Increasing TMS1 polypeptide level to an above normal level. The CpG
island region of TMS1 or its fragments are used to study the methylation
patterns apart from any coding region contained in it.
The present sequence is human target of methylation-induced silencing-1
(TMS1) protein.
Sequence 195 AA;
Query Match 100.0%; Score 990; DB 22; Length 195;
Best Local Similarity 100.0%; Pred. No. 3.2e-101;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLDTKLVSPY 60
DB 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLDTKLVSPY 60
QY 61 LETYGAELTANVLRLDMGLQEMAGLQAAHQSGGAAPAGIQAAPPQSAAKPGLHFIDQHRA 120
DB 61 LETYGAELTANVLRLDMGLQEMAGLQAAHQSGGAAPAGIQAAPPQSAAKPGLHFIDQHRA 120
QY 121 ALIARTVNTVEWLLDALYGVKLTDEQYQAVRAEPTNPSKMKLPSFTPAWNTCKDILLQA 180
DB 121 ALIARTVNTVEWLLDALYGVKLTDEQYQAVRAEPTNPSKMKLPSFTPAWNTCKDILLQA 180
QY 181 LRESQSYLVEDLERS 195
DB 181 LRESQSYLVEDLERS 195

RESULT 4
AAB20085
ID AAB20085 standard; Protein; 195 AA.
XX
AC AAB20085;
XX
DT 23-APR-2001 (first entry)
XX Human CARD-5 protein.
DE
XX CARD-5; caspase recruitment domain; human; cancer; infection;
KW autoimmune disease; neurological disease; haematological disease;
KW immune disease; inflammation; antitumour; antiseptic;
KW immunomodulator; antiinflammatory; apoptosis; diagnosis;
KW gene therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 111..181
FT /note= "CARD"
XX
EN WO200100826-A2.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-US17691.
XX
PR 28-JUN-1999; 99US-0340620.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Bertin J;
XX
DR WPI; 2001-061973/07.
DR N-PSDB; AAF30007.
XX
PT Isolated intracellular proteins predicted to be involved in regulating
PT caspase activation are used for diagnosis and treatment of e.g. cancer,
PT viral infections, autoimmune diseases, neurological diseases and
PT haematological disorders -
XX
PS Claim 9; Fig 21; 208pp; English.
XX
CC The present sequence is that of human caspase recruitment domain 5
CC (CARD-5), an intracellular protein predicted to be involved in
CC regulating caspase activation. The sequence is predicted from an
CC isolated cDNA clone (see AAF30007). Methods of diagnosing and
CC treating patients suffering from a disorder associated with an
CC abnormal level or rate of apoptotic cell death, abnormal activity
CC of the Fas/APO-1 receptor complex, abnormal activity of the tumour
CC necrosis factor receptor complex or abnormal activity of a caspase
CC involve administering a compound that modulates the expression or
CC activity of CARD-3, CARD-4, CARD-5 or CARD-6 e.g. using gene
CC therapy methods. Such disorders include cancer, viral infection,
CC autoimmune disorders, neurological diseases, haematological
CC disorders, inflammatory disorders and immune disorders. CARD-3,
CC -4, -5 and -6 proteins can be used to regulate cell proliferation,
CC cell survival and cell growth. They can also be used to screen
CC drugs or compounds that modulate their activity or expression and
CC to treat disorders associated with insufficient or excessive
CC production of CARD-3, -4, -5 or -6 protein, or production of an
CC aberrant protein.
XX
SQ Sequence 195 AA;
Query Match 100.0%; Score 990; DB 22; Length 195;
Best Local Similarity 100.0%; Pred. No. 3.2e-101;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLDTKLVSPY 60
DB 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLDTKLVSPY 60

QY 61 LETYGAELTANVLRDMGLQEMAGLOQAATHOGSGAAPAGIQAPPOSAAKPGLHFDQHRA 120
 DB 61 LETYGAELTANVLRDMGLQEMAGLOQAATHOGSGAAPAGIQAPPOSAAKPGLHFDQHRA 120
 QY 121 ALIARTVNVWLLDALYGVLTDEQYQAVRAEPTNPSKMKLFSTPPAWNWTCKDLLQA 180
 DB 121 ALIARTVNVWLLDALYGVLTDEQYQAVRAEPTNPSKMKLFSTPPAWNWTCKDLLQA 180
 QY 181 LRESQSYLVEDLERS 195
 DB 181 LRESQSYLVEDLERS 195

RESULT 5
 AAU99353
 ID AAU99353 standard; Protein; 195 AA.
 AC AAU99353;
 XX
 XX
 DT 07-OCT-2002 (first entry)
 XX
 DE Human caspase recruitment domain-5 (CARD-5) protein.
 KW Human; caspase recruitment domain-5; CARD-5; antiinflammatory;
 KW immunosuppressive; caspase; cysteinyl aspartate-specific proteinase;
 KW apoptosis; nuclear factor-kappa B; NF-kappaB; transcription factor;
 KW cell proliferation; gene therapy; immune disorder;
 KW chronic inflammatory disease; Hashimoto's thyroiditis; graft rejection;
 KW sarcoidosis; atopy; asthma; allergy; glomerular nephritis;
 KW human immunodeficiency virus; HIV; bacterial infection; tuberculosis;
 KW lepromatous leprosy; autoimmune disorder; systemic lupus erythematosus;
 KW arthritis; cell depletion; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; spinal muscular atrophy; haematologic disease;
 KW myelodysplastic syndrome; aplastic anaemia; myocardial infarction;
 KW stroke.
 XX
 XX Homo sapiens.
 XX
 XX
 .FH Key Location/Qualifiers
 FT Domain 111..195
 FT /label= CARD_domain
 XX
 PN W0200244354-A2.
 XX
 PD 06-JUN-2002.
 XX
 XX 29-NOV-2001; 2001WO-US44894.
 PF
 PR 01-DEC-2000; 2000US-0728721.
 PR 24-APR-2001; 2001US-0841879.
 XX
 XX (MILL-) MILLENIUM PHARM INC.
 PA
 PI Bertin J;
 PI
 DR WPI; 2002-557538/59.
 DR N-PSDB; ABR87966.
 XX
 XX Novel isolated murine or human caspase recruitment domain (CARD)-5
 PT polypeptide, useful for treating immune disorders such as Hashimoto's
 PT thyroiditis, graft rejection, allergy, glomerular nephritis,
 PT tuberculosis -
 XX
 XX Claim 22; Fig 3; 100pp; English.
 PS
 XX
 XX The invention discloses the isolated polypeptides, and encoding nucleic
 CC acids, of murine and human caspase recruitment domain (CARD)-5. Caspases
 CC (cysteinyl aspartate-specific proteinases) are central to the apoptotic
 CC program and responsible for the degradation of cellular proteins that
 CC lead to the morphological changes seen in cells undergoing apoptosis.
 CC Caspases interact with other caspases via their CARDS and different
 CC subtypes of CARDS may confer binding specificity. CARD-5 is an

CC intracellular protein that is predicted to be involved in regulating
 CC caspase activation. CARD-5 activates the nuclear factor-kappa B
 CC (NF-kappaB) transcription factor pathway and binds the CARDS of
 CC caspase-1, CARD-7 and itself. CARD-5 can, therefore, modulate CARD-5
 CC activity and NF-kappaB activation, regulate cell growth and cell death
 CC and be used in gene therapy. The CARD-5 polypeptides are useful for
 CC identifying compounds which bind to them and modulate their activity and
 CC for detecting the presence of CARD-5 in a sample. CARD-5 polypeptides,
 CC nucleic acids, antibodies and modulators of CARD-5 expression or activity
 CC can be used to treat immune disorders such as chronic inflammatory
 CC diseases and disorders, Hashimoto's thyroiditis, graft rejection,
 CC sarcoidosis, atopic conditions (such as asthma and allergy), glomerular
 CC nephritis, human immunodeficiency virus (HIV) and bacterial infections
 CC (including tuberculosis and lepromatous leprosy) and in screening and
 CC detection assays. Modulators of CARD-5 activity or expression are also
 CC useful for treating autoimmune disorders, such as systemic lupus
 CC erythematosus and arthritis, cell depletion, neurological disorders,
 CC such as Alzheimer's disease, Parkinson's disease and spinal muscular
 CC atrophy, haematologic diseases, such as myelodysplastic syndrome and
 CC aplastic anaemia, myocardial infarction and stroke. The sequence
 CC presented is the human caspase recruitment domain-5 (CARD-5) protein.
 XX
 SQ Sequence 195 AA;
 Query Match 100.0%; Score 990; DB 23; Length 195;
 Best Local Similarity 100.0%; Pred. No. 3.2e-101;
 Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGRARDAILDALENLTABELKKFKLLSVLRGGRIPRGALLSMDALDITDKLVSFY 60
 DB 1 MGRARDAILDALENLTABELKKFKLLSVLRGGRIPRGALLSMDALDITDKLVSFY 60
 QY 61 LETYGAELTANVLRDMGLQEMAGLOQAATHOGSGAAPAGIQAPPOSAAKPGLHFDQHRA 120
 DB 61 LETYGAELTANVLRDMGLQEMAGLOQAATHOGSGAAPAGIQAPPOSAAKPGLHFDQHRA 120
 QY 121 ALIARTVNVWLLDALYGVLTDEQYQAVRAEPTNPSKMKLFSTPPAWNWTCKDLLQA 180
 DB 121 ALIARTVNVWLLDALYGVLTDEQYQAVRAEPTNPSKMKLFSTPPAWNWTCKDLLQA 180
 QY 181 LRESQSYLVEDLERS 195
 DB 181 LRESQSYLVEDLERS 195

RESULT 6
 AA017854
 ID AA017854 standard; Protein; 195 AA.
 XX
 AC AA017854;
 XX
 DT 20-AUG-2002 (first entry)
 XX
 DE Pyrin domain containing protein Pycard.
 XX
 KW Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
 KW antiarteriosclerotic; antipsoaric; antibacterial; virucide;
 KW neuroprotective; antiarthritic; antirheumatic; antiasthmatic;
 KW nephrotropic; osteopathic; nootropic; intracellular signal transduction;
 KW inflammation; Alzheimer's disease; infection; psoriasis; asthma;
 KW arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
 KW osteoarthritis; glomerulonephritis.
 XX
 OS Unidentified.
 XX
 PN W0200240668-A2.
 PD 23-MAY-2002.
 XX
 XX 30-OCT-2001; 2001WO-EP12545.
 PF
 XX 15-NOV-2000; 2000DE-1056687.
 PR
 PR 30-NOV-2000; 2000DE-1059595.

XX (APOT-) APOTEC RES & DEV LTD.
 XX Tschoep J, Martinon F;
 PI WPI, 2002-427093/45.
 XX N-PSDB; AAL47126.
 DR New DNA encoding protein with pyrin domain, useful for treating
 XX diseases involving impaired signal transduction, particularly
 PT inflammation, also proteins and antibodies -
 PT
 XX Claim 5; Fig 1; 116pp; German.
 PS
 XX The present invention relates the DNA and their encoded proteins, where
 CC the proteins contain at least one PYD (pyrin) domain. These can be used
 CC to treat diseases associated with impaired intracellular signal
 CC transduction, particularly inflammation such as psoriasis,
 CC arteriosclerosis, bacterial or viral infections (particularly meningitis
 CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
 CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
 CC and Parkinson's diseases. The present sequence is a protein of the
 CC invention.
 XX
 XX Sequence 195 AA;
 SQ
 Query Match 100.0%; Score 990; DB 23; Length 195;
 Best Local Similarity 100.0%; Pred. No. 3.2e-101;
 Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGRARDAILDALENLTAEELKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60
 DB 1 MGRARDAILDALENLTAEELKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60
 QY 61 LETYGAELTANVLRDMGLQEMAGLOQAATHQSGGAAPAGIQAPPOSAAKPGLHFDQHRA 120
 DB 61 LETYGAELTANVLRDMGLQEMAGLOQAATHQSGGAAPAGIQAPPOSAAKPGLHFDQHRA 120
 QY 121 ALIARTVNVVWLLDALYGVKLTDEQYQAVRAEPTNPSKMKLFSTPAWNTCKDLLQA 180
 DB 121 ALIARTVNVVWLLDALYGVKLTDEQYQAVRAEPTNPSKMKLFSTPAWNTCKDLLQA 180
 QY 181 LRESQSYLVEDLERS 195
 DB 181 LRESQSYLVEDLERS 195
 RESULT 7
 ABG71635
 ID ABG71635 standard; Protein; 195 AA.
 AC
 XX ABG71635;
 XX
 XX 10-JAN-2003 (first entry)
 DT
 XX Human caspase recruitment domain-5 (CARD-5).
 DE
 XX Human; caspase activity; caspase recruitment domain-5; CARD-5;
 KW caspase-1; pseudo-interleukin-1 beta converting enzyme; IL-1beta;
 KW pseudolICE; ICEBERG; cell growth; cell death; inflammation; CARD-7;
 KW apoptosis; caspase activation; cancer; follicular lymphoma;
 KW leukaemia; melanoma; colon cancer; lung carcinoma; viral infection;
 KW autoimmune disease; systemic lupus erythematosus; reactive arthritis;
 KW human immunodeficiency virus infection; HIV infection; AIDS;
 KW Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
 KW myocardial infarction; stroke; inflammatory disorder; Crohn's disease;
 KW insulin dependent diabetes mellitus; multiple sclerosis; psoriasis;
 KW graft rejection; allergic rhinitis; food allergy; conjunctivitis;
 KW glomerular nephritis; cytostatic; virucide; immunosuppressive;
 KW dermatological; nephrotropic; neuroprotective; cardiant;
 KW caspase recruitment domain-7.
 XX
 XX Homo sapiens.

XX US2002128198-A1.
 XX 12-SEP-2002.
 XX 27-NOV-2001; 2001US-0996617.
 XX 28-JUN-1999; 99US-0340620.
 PR 27-OCT-1999; 99US-0428252.
 PR 15-AUG-2001; 2001US-0931071.
 XX (BERT/) BERTIN J.
 PA Bertin J;
 XX WPI; 2003-028967/02.
 DR N-PSDB; ABS56032.
 XX Identifying modulator of CARD-7 and CARD-5 interaction, by contacting
 PT CARD-7 and CARD-5 in presence of test compound, measuring their
 PT binding, and identifying modulator, when binding of CARD-7 to CARD-5 is
 PT altered -
 XX Disclosure; Fig 7; 43pp; English.
 PS
 XX The present invention relates to methods of identifying compounds
 CC that regulate caspase activity using caspase recruitment domain-7
 CC (CARD-7) or caspase recruitment domain-8 (CARD-8). In particular, a
 CC method for identifying a compound that modulates the interaction
 CC between CARD-8 and caspase-1, pseudo-interleukin-1 (IL-1) beta
 CC converting enzyme (pseudolICE) or ICEBERG is disclosed. CARD-7 and
 CC CARD-8 molecules are useful as modulating agents in regulating a
 CC variety of cellular processes including cell growth, cell death, and
 CC inflammation. The methods of the invention are useful for identifying
 CC compounds that have the ability to increase/decrease apoptosis, or
 CC comprise the ability to induce caspase activation. The methods are
 CC useful for treating a disorder associated with inappropriate apoptosis
 CC or inappropriate inflammation. The methods are useful for treating
 CC disorders associated with an undesirably low rate of apoptosis such
 CC as cancer (preferably follicular lymphoma, chronic myelogenous
 CC leukaemia, melanoma, colon cancer, lung carcinoma, etc), viral
 CC infections, autoimmune diseases caused by low levels of apoptosis
 CC (e.g. systemic lupus erythematosus, immune-mediated glomerulonephritis,
 CC and arthritis). The methods are also useful for treating disorders with
 CC undesirably high rates of apoptosis such as human immunodeficiency
 CC virus (HIV) infection, Alzheimer's disease, Parkinson's disease,
 CC amyotrophic lateral sclerosis (ALS), retinitis pigmentosa, spinal
 CC muscular atrophy, various forms of cerebellar degeneration, anaemia
 CC associated with chronic disease, aplastic anaemia, chronic neutropaenia,
 CC myelodysplastic syndromes, myocardial infarction, stroke, and
 CC various inflammatory disorders (e.g. Crohn's disease, reactive
 CC arthritis, insulin dependent diabetes mellitus, multiple sclerosis,
 CC psoriasis, graft rejection, allergic rhinitis, food allergies,
 CC conjunctivitis, glomerular nephritis, etc). The present sequence
 CC represents human CARD-5. The sequence encoding CARD-5 is used to
 CC identify cDNA encoding CARD-7.
 XX
 XX Sequence 195 AA;
 SQ
 Query Match 100.0%; Score 990; DB 24; Length 195;
 Best Local Similarity 100.0%; Pred. No. 3.2e-101;
 Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGRARDAILDALENLTAEELKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60
 DB 1 MGRARDAILDALENLTAEELKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60
 QY 61 LETYGAELTANVLRDMGLQEMAGLOQAATHQSGGAAPAGIQAPPOSAAKPGLHFDQHRA 120
 DB 61 LETYGAELTANVLRDMGLQEMAGLOQAATHQSGGAAPAGIQAPPOSAAKPGLHFDQHRA 120
 QY 121 ALIARTVNVVWLLDALYGVKLTDEQYQAVRAEPTNPSKMKLFSTPAWNTCKDLLQA 180
 DB 121 ALIARTVNVVWLLDALYGVKLTDEQYQAVRAEPTNPSKMKLFSTPAWNTCKDLLQA 180

Db 121 ALIARVTNVEWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSEFTPAWNWTKOLLQA 180
QY 181 LRESQSYLVEDLERS 195
Db 181 LRESQSYLVEDLERS 195
RESULT 8
ID AAG74647
AC AAG74647;
DT 03-SEP-2001 (first entry)
DE Human colon cancer antigen protein SEQ ID NO:5411.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma.
OS Homo sapiens.
XX
PN WO20012920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26524.
XX
PR 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
DR WPI; 2001-235357/24.
DR N-PSDB; AAH34052.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX
PS Claim 11; Page 7035-7036; 9803pp; English.
XX
CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing
CC inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ Sequence 205 AA;
Query Match 100.0%; Score 990; DB 22; Length 205;
Best Local Similarity 100.0%; Pred. No. 3.4e-101;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRARDAILDALENLTAEELKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60
Db 11 MGRARDAILDALENLTAEELKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 70

QY 61 LETYGAEITANVLRDMGLQEMAGQLQAATHQSGGAAPAGIQAPPOSAAKPGLHFTDOHRA 120
Db 71 LETYGAEITANVLRDMGLQEMAGQLQAATHQSGGAAPAGIQAPPOSAAKPGLHFTDOHRA 130
QY 121 ALIARVTNVEWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSEFTPAWNWTKOLLQA 180
Db 131 ALIARVTNVEWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSEFTPAWNWTKOLLQA 190
QY 181 LRESQSYLVEDLERS 195
Db 191 LRESQSYLVEDLERS 205
RESULT 9
ID AAE00594 standard; Protein; 176 AA.
XX
AC AAE00594;
DT 02-JUL-2001 (first entry)
DE Alternatively spliced form of human TMS1 protein (lacking exon2).
XX
KW Human; target of methylation-induced silencing-1; TMS1; cytostatic;
KW antiproliferative; apoptosis inducer; gene therapy; CpG island;
KW caspase-recruiting domain; CARD; cancer; breast.
XX
OS Homo sapiens.
XX
PN WO200129235-A2.
XX
PD 26-APR-2001.
XX
PF 18-OCT-2000; 2000WO-US28747.
XX
PR 18-OCT-1999; 99US-0159975.
XX
PA (UYEM-) UNIV EMORY.
XX
PI Vertino PM;
XX
DR WPI; 2001-290922/30.
DR N-PSDB; AAD03906.
XX
CC Novel gene TMS1, transcriptionally silenced due to increased
CC methylation useful for identifying subject at risk of developing tumor
CC characterized by abnormal methylation, for treating cancer by inducing
CC apoptosis -
XX
PS Claim 85; Page 123; 124pp; English.
XX
CC The invention relates to identification of target of methylation-induced
CC silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to
CC abnormal methylation of a CpG island in its 5' regulatory region. TMS1
CC consists of a carboxy terminal caspase-recruiting domain (CARD) and
CC plays a role in induction of apoptosis. TMS1 gene and protein are useful
CC as tools for diagnosing and treating a subject at risk of developing
CC cancer (e.g. breast cancer) characterised by abnormal CpG methylation or
CC abnormally low levels of TMS1 expression products. Unique fragments of
CC TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.
CC TMS1 molecule is also useful for treating abnormal cell proliferation by
CC increasing TMS1 polypeptide level to an above normal level. The CpG
CC island region of TMS1 or its fragments are used to study the methylation
CC patterns apart from any coding region contained in it.
CC The present sequence is alternatively applied form of human target of
CC methylation-induced silencing-1 (TMS1) protein lacking exon2.
XX
SQ Sequence 176 AA;
Query Match 88.2%; Score 873.5; DB 22; Length 176;
Best Local Similarity 90.3%; Pred. No. 2.3e-88;
Matches 176; Conservative 0; Mismatches 0; Indels 19; Gaps 1;

QY 1 MGRARDAILDALENLTAEELKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSFY 60
 Db 1 MGRARDAILDALENLTAEELKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSFY 60
 QY 61 LETYGAELTANVLRDMGLQEMAGOLQAATHQSGGAAPAGIQAPPOSAAKPGHLHFIDQHRA 120
 Db 61 LETYGAELTANVLRDMGLQEMAGOLQAATHQSGGAAPAGIQAPPOSAAKPGHLHFIDQHRA 101
 QY 121 ALIARVTNVVWLLDALYKGLVLTDEQYQAVRAEPTNPSKMRKLFSTPAWNWTKDOLLQA 180
 Db 102 ALIARVTNVVWLLDALYKGLVLTDEQYQAVRAEPTNPSKMRKLFSTPAWNWTKDOLLQA 161
 QY 181 LRESQSYLVEDLERS 195
 Db 162 LRESQSYLVEDLERS 176

RESULT 10

AAB43675

ID AAB43675 standard; Protein; 190 AA.

AC AAB43675;

DT 08-FEB-2001 (first entry)

DE Human cancer associated protein sequence SEQ ID NO:1120.

XX Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
 KW antidiabetic; antiaschmatic; antirheumatic; antithrombotic; antiviral;
 KW antinflammatory; antihypertensive; antiallergic; antibacterial; cardiant;
 KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening.

OS Homo sapiens.

XX WO200005350-A1.

PN 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05882.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2000-587533/55.

DR N-PSDB; AAC77884.

XX Novel isolated nucleic acids comprising sequences encoding peptides
 useful for treating or diagnosing e.g. cancer -

PS Claim 11; Page 1731-1732; 2352pp; English.

CC AAC77607 to AAC78448 encode the human cancer associated proteins given
 in AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnery; immunomodulator;
 CC antidiabetic; antiaschmatic; antirheumatic; antithrombotic;
 CC antinflammatory; antihypertensive; antiallergic; antibacterial; antiviral;
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
 CC nootropic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of

CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention.

SQ Sequence 190 AA;

Query Match 76.7%; Score 759; DB 21; Length 190;

Best Local Similarity 95.6%; Pred. No. 1.2e-75;

Matches 153; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 1 MGRARDAILDALENLTAEELKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSFY 60

Db 11 MGRARDAILDALENLTAEELKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSFY 70

QY 61 LETYGAELTANVLRDMGLQEMAGOLQAATHQSGGAAPAGIQAPPOSAAKPGHLHFIDQHRA 120

Db 71 LETYGAELTANVLRDMGLQEMAGOLQAATHQSGGAAPAGIQAPPOSAAKPGHLHFIDQHRA 130

QY 121 ALIARVTNVVWLLDALYKGLVLTDEQYQAVRAEPTNPSKMR 160

Db 131 ALIARVTNVVWLLDALYKGLVLTDEQYQAVR--PSPPTQAR 168

RESULT 11

AAB00592

ID AAB00592 standard; Protein; 193 AA.

XX AAB00592;

XX 02-JUL-2001 (first entry)

DE Mouse target of methylation-induced silencing-1 (TMS1) protein.

XX Mouse; target of methylation-induced silencing-1; TMS1; cytostatic;
 KW antiproliferative; apoptosis inducer; gene therapy; CpG island;
 KW caspase-recruiting domain; CARD; cancer; breast.

XX Mus musculus.

XX WO200129235-A2.

XX 26-APR-2001.

XX 18-OCT-2000; 2000WO-US28747.

XX 18-OCT-1999; 99US-0159975.

XX (UYEM-) UNIV EMORY.

XX Vertino PM;

XX WPI; 2001-290922/30.

XX N-PSDB; AAB03904.

XX Novel gene TMS1, transcriptionally silenced due to increased

PT methylation useful for identifying subject at risk of developing tumor
 PT characterized by abnormal methylation, for treating cancer by inducing
 PT apoptosis -

PS Claim 85; Page 120; 124pp; English.

XX The invention relates to identification of target of methylation-induced
 CC silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to
 CC abnormal methylation of a CpG island in its 5' regulatory region. TMS1
 CC consists of a carboxy terminal caspase-recruiting domain (CARD) and
 CC plays a role in induction of apoptosis. TMS1 gene and protein are useful
 CC as tools for diagnosing and treating a subject at risk of developing
 CC cancer (e.g. breast cancer) characterised by abnormal CpG methylation or

CC abnormally low levels of TMS1 expression products. Unique fragments of
 CC TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.
 CC TMS1 molecule is also useful for treating abnormal cell proliferation by
 CC increasing TMS1 polypeptide level to an above normal level. The CpG
 CC island region of TMS1 or its fragments are used to study the methylation
 CC patterns apart from any coding region contained in it.
 CC The present sequence is mouse target of methylation-induced silencing-1
 CC (TMS1) protein.
 XX Sequence 193 AA;
 SQ

Query Match 69.1%; Score 684; DB 22; Length 193;
 Best Local Similarity 71.8%; Pred. No. 2.6e-67;
 Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;

QY 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60
 DB 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60
 QY 61 LETYGAELTANVLRDMGLQEMAGLOQAATHOGSGAAPAGIQAPPOSAAKPGHLFDQHRA 120
 DB 61 LESYGLELTMTVLRDMGLQELAEQLQ-TTKEESGAVAAAASVPAQSTARTG-HFVDQHRQ 118
 QY 121 ALIARVTNVVWLLDALYGVLTDEQYQAVRAEPTNPSKMKLFSFTPAWNTCKDILLQA 180
 DB 119 ALIARVTEVDGVLDAHGSVLTEGQYQAVRAETTSQDKMKLFSFVPSNNLTCKDILLQA 178
 QY 181 LRESQSYLVEDLERS 195
 DB 179 LKEIHPYLVMDLEQS 193

RESULT 12
 AAB20086
 ID AAB20086 standard; Protein; 193 AA.
 AC AAB20086;
 DT 23-APR-2001 (first entry)
 DE Mouse CARD-5 protein.
 KW CARD-5; caspase recruitment domain; mouse; cancer; infection;
 KW autoimmune disease; neurological disease; haematological disease;
 KW immune disease; inflammation; anticancer; antiseptic;
 KW immunomodulator; antiinflammatory; apoptosis; diagnosis;
 KW gene therapy.
 OS Mus sp.
 XX

Key Location/Qualifiers
 FT Domain 110..179
 FT /note= "CARD"
 PN WO200100826-A2.
 XX
 PD 04-JAN-2001.
 XX
 PF 28-JUN-2000; 2000WO-US17691.
 XX
 PR 28-JUN-1999; 99US-0340620.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Bertin J;
 XX
 DR WPI; 2001-061973/07.
 DR N-PSDB; AAF30008.
 XX

Isolated intracellular proteins predicted to be involved in regulating
 PT caspase activation are used for diagnosis and treatment of e.g. cancer,
 PT viral infections, autoimmune diseases, neurological diseases and
 PT haematological disorders -

XX Claim 9; Fig 19; 20pp; English.
 PS
 CC The present sequence is that of mouse caspase recruitment domain 5
 CC (CARD-5), an intracellular protein predicted to be involved in
 CC regulating caspase activation. The sequence is predicted from an
 CC isolated cDNA clone (see AAF30008). It shows 71.8% amino acid
 CC identity to human CARD-5 (see AAB20085). Methods of diagnosing
 CC and treating patients suffering from a disorder associated with an
 CC abnormal level or rate of apoptotic cell death, abnormal activity
 CC of the Fas/APO-1 receptor complex, abnormal activity of the tumour
 CC necrosis factor receptor complex or abnormal activity of a caspase
 CC involve administering a compound that modulates the expression or
 CC activity of CARD-3, CARD-4, CARD-5 or CARD-6 e.g. using gene
 CC therapy methods. Such disorders include cancer, viral infection,
 CC autoimmune disorders, neurological diseases, haematological
 CC disorders, inflammatory disorders and immune disorders. CARD-3,
 CC -4, -5 and -6 proteins can be used to regulate cell proliferation,
 CC cell survival and cell growth. They can also be used to screen
 CC drugs or compounds that modulate their activity or expression and
 CC to treat disorders associated with insufficient or excessive
 CC production of CARD-3, -4, -5 or -6 protein, or production of an
 CC aberrant protein.
 XX Sequence 193 AA;
 SQ

Query Match 69.1%; Score 684; DB 22; Length 193;
 Best Local Similarity 71.8%; Pred. No. 2.6e-67;
 Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;

QY 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60
 DB 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60
 QY 61 LETYGAELTANVLRDMGLQEMAGLOQAATHOGSGAAPAGIQAPPOSAAKPGHLFDQHRA 120
 DB 61 LESYGLELTMTVLRDMGLQELAEQLQ-TTKEESGAVAAAASVPAQSTARTG-HFVDQHRQ 118
 QY 121 ALIARVTNVVWLLDALYGVLTDEQYQAVRAEPTNPSKMKLFSFTPAWNTCKDILLQA 180
 DB 119 ALIARVTEVDGVLDAHGSVLTEGQYQAVRAETTSQDKMKLFSFVPSNNLTCKDILLQA 178
 QY 181 LRESQSYLVEDLERS 195
 DB 179 LKEIHPYLVMDLEQS 193

RESULT 13
 AAU99352
 ID AAU99352 standard; Protein; 193 AA.
 AC AAU99352;
 DT 07-OCT-2002 (first entry)
 DE Mouse caspase recruitment domain-5 (CARD-5) protein.
 XX
 KW Mouse; caspase recruitment domain-5; CARD-5; antiinflammatory;
 KW immunosuppressive; caspase; cysteinyl aspartate-specific proteinase;
 KW apoptosis; nuclear factor-kappa B; NF-kappaB; transcription factor;
 KW cell proliferation; gene therapy; immune disorder;
 KW chronic inflammatory disease; Hashimoto's thyroiditis; graft rejection;
 KW sarcoidosis; atopy; asthma; allergy; glomerular nephritis;
 KW human immunodeficiency virus; HIV; bacterial infection; tuberculosis;
 KW lepromatous leprosy; autoimmune disorder; systemic lupus erythematosus;
 KW arthritis; cell depletion; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; spinal muscular atrophy; haematologic disease;
 KW myelodysplastic syndrome; aplastic anaemia; myocardial infarction;
 KW stroke.
 OS Mus sp.
 XX
 XX Key Location/Qualifiers
 FH

Domain 110..193
/label= CARD_domain
WO200244354-A2.
06-JUN-2002.
29-NOV-2001; 2001WO-US44894.
01-DEC-2000; 2000US-0728721.
24-APR-2001; 2001US-0841879.
(MILL-) MILLENIUM PHARM INC.
Bertin J;
WPI: 2002-557538/59.
N-PSDB; ABR87964.
Novel isolated murine or human caspase recruitment domain (CARD)-5 polypeptide, useful for treating immune disorders such as Hashimoto's thyroiditis, graft rejection, allergy, glomerular nephritis, tuberculosis -
Claim 22; Fig 1; 100pp; English.
The invention discloses the isolated polypeptides, and encoding nucleic acids, of murine and human caspase recruitment domain (CARD)-5. Caspases (cysteine aspartate-specific proteinases) are central to the apoptotic program and responsible for the degradation of cellular proteins that lead to the morphological changes seen in cells undergoing apoptosis. Caspases interact with other caspases via their CARDS and different subtypes of CARDS may confer binding specificity. CARD-5 is an intracellular protein that is predicted to be involved in regulating caspase activation. CARD-5 activates the nuclear factor-kappa B (NF-kappaB) transcription factor pathway and binds the CARDS of caspase-1, CARD-7 and itself. CARD-5 can, therefore, modulate CARD-5 activity and NF-kappaB activation, regulate cell growth and cell death and be used in gene therapy. The CARD-5 polypeptides are useful for identifying compounds which bind to them and modulate their activity and for detecting the presence of CARD-5 in a sample. CARD-5 polypeptides, nucleic acids, antibodies and modulators of CARD-5 expression or activity can be used to treat immune disorders such as chronic inflammatory diseases and disorders, Hashimoto's thyroiditis, graft rejection, sarcoidosis, atopic conditions (such as asthma and allergy), glomerular nephritis, human immunodeficiency virus (HIV) and bacterial infections (including tuberculosis and lepromatous leprosy) and in screening and detection assays. Modulators of CARD-5 activity or expression are also useful for treating autoimmune disorders, such as systemic lupus erythematosus and arthritis, cell depletion, neurological disorders, such as Alzheimer's disease, Parkinson's disease and spinal muscular atrophy, haematologic diseases, such as myelodysplastic syndrome and aplastic anaemia, myocardial infarction and stroke. The sequence presented is the mouse caspase recruitment domain-5 (CARD-5) protein.
Sequence 193 AA;
Query Match 69.1%; Score 684; DB 23; Length 193;
Best Local Similarity 71.8%; Pred. No. 2.6e-67;
Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;
QY 1 MGRARDAILDALENTAEELKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60
DB 1 MGRARDAILDALENLSGDELKFKMLLTQVREGYGRIPRGALLQMDALDITDKLVSY 60
QY 61 LETYGAELTANVLNRMGLQEMAGLQQAATHQSGAGAPAGIOAPPOSAKPGFLHFDQHRA 120
DB 61 LESYGLELTMTVLNRMGLQELAEQLQ-TTKEESGAVAAASVPAQSTARTG-HFVDQHRQ 118
QY 121 ALIARTVNEWLLDALYGVKLTDEQVAVRAEPTNPSKMKLFSPFPANWTCKOLLLOA 180
DB 119 ALIARTVTEVDGLDALHGSVLTEGQYAVRAETTSQDKMKLFSPFVPSWNLCKDSSLQA 178

QY 181 LRSQSILVEDLERS 195
DB 179 LKEIHPVLVMDLEQS 193
RESULT 14
AAE00593
ID AAE00593 standard; Protein; 171 AA.
XX AAE00593;
AC AAE00593;
XX 02-JUL-2001 (first entry)
DT Rat target of methylation-induced silencing-1 (TMS1) partial protein.
DE Rat; target of methylation-induced silencing-1; TMS1; cytostatic;
XX antiproliferative; apoptosis inducer; gene therapy; CpG island;
KW caspase-recruiting domain; CARD; cancer; breast.
XX
OS Rattus norvegicus.
PH Key Location/Qualifiers
FT Misc-difference 36 /note= "Encoded by AGY"
FT
XX WO200129235-A2.
XX 26-APR-2001.
XX 18-OCT-2000; 2000WO-US28747.
XX 18-OCT-1999; 99US-0159975.
PA (UYEM-) UNIV EMORY.
XX Vertino PM;
PI WPI: 2001-290922/30.
XX N-PSDB; AAD03905.
XX Novel gene TMS1, transcriptionally silenced due to increased methylation useful for identifying subject at risk of developing tumor characterized by abnormal methylation, for treating cancer by inducing apoptosis -
PS Claim 85; Page 121; 124pp; English.
XX The invention relates to identification of target of methylation-induced silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to abnormal methylation of a CpG island in its 5' regulatory region. TMS1 consists of a carboxy terminal caspase-recruiting domain (CARD) and plays a role in induction of apoptosis. TMS1 gene and protein are useful as tools for diagnosing and treating a subject at risk of developing cancer (e.g. breast cancer) characterised by abnormal CpG methylation or abnormally low levels of TMS1 expression products. Unique fragments of TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.
XX TMS1 molecule is also useful for treating abnormal cell proliferation by increasing TMS1 polypeptide level to an above normal level. The CpG island region of TMS1 or its fragments are used to study the methylation patterns apart from any coding region contained in it.
XX The present sequence is rat target of methylation-induced silencing-1 (TMS1) partial protein.
XX Sequence 171 AA;
Query Match 58.2%; Score 576; DB 22; Length 171;
Best Local Similarity 67.1%; Pred. No. 2e-55;
Matches 116; Conservative 18; Mismatches 37; Indels 2; Gaps 2;
QY 23 PKLLLSVPLREGYGRIPRGALLSMDALDITDKLVSVFYETYGAEITANVLNRMGLQEMA 82
DB 1 FKIKLLTAPVREGYGRIPRGALLQMDPDLTDLKLVXYLEGYGLTMTVLNRMGLQELA 60

QY 83 GQLQAATHOGSGAAPAGIQAPPOSAAKPGLHFIDQHRAALIARVTNVWLLDALYGVLT 142
Db 61 EQQKIMEE-SGAVATATSPAGQTARTE-HFVDQHRQALIARVTEVDGLLDALYGVLT 118
QY 143 DEQYQAVRAEPTNPSSKMRKLFSTPAWNWTKDOLLQALRESQSYLVEIDERS 195
Db 119 EGOYQAVRAEPTTNQNRKLFSPAPAWNLTCKNLFLEALRQTPQYLVLTLEQS 171

RESULT 15
AAG74648
ID AAG74648 standard; Protein; 158 AA.
XX
AC AAG74648;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen protein SEQ ID NO:5412.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX colorectal carcinoma.
OS Homo sapiens.
XX
PN WO200122920-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26524.
XX
PR 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
WPI; 2001-235357/24.
DR N-PSDB; AAH34053.
XX
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX
PS Claim 11; Page 7036-7037; 9803pp; English.
XX
CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where
CC the proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate P
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of P by expressing P.
CC Inactive proteins or to supplement the patients own production of P.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAB77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ Sequence 158 AA;

Query Match 57.0%; Score 564.5; DB 22; Length 158;
Best Local Similarity 75.8%; Pred. No. 3.4e-54;
Matches 119; Conservative 3; Mismatches 2; Indels 33; Gaps 2;
QY 4 ARDAILDALLENLTABELKKFLKLLSVPLREGYIPRGALLSMDALDITDKLVSFYLET 63
|||||

Db 13 ARDAILDALLENLTABELKKF-----KLVSFYLET 41
QY 64 YGAELTANVLKRDMLQEWAGOLQAATHOGSGAAPAGIQAPPOSAAKPGLHFIDQHRAALI 123
Db 42 YGAELTANVLKRDMLQEWAGOLQAATHOGSGAAPAGIQAPPOSAAKPGLHFIDQHRAALI 101
QY 124 ARVTNVEWLLDALYGVLTDEQYQAVRAEPTNPSSKMR 160
Db 102 ARVTNVEWLLDALYGVLTDEQYQAVR--PSPTQAR 136

Search completed: January 29, 2004, 13:45:54
Job time : 9.48584 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 13:44:42 ; Search time 4.80296 Seconds
(without alignments)
3904.448 Million cell updates/sec

Title: US-09-996-617-8
Perfect score: 990
Sequence: 1 MGRARDALDALENLTAEEL.....LLQLALRSQSIVLVEDLRS 195

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	250	25.3	1192	T17255	hypothetical prote
2	87	8.8	323	T49723	hypothetical prote
3	86.5	8.7	372	S23326	Gene ML2.2 protein
4	86.5	8.7	439	S75545	hypothetical prote
5	85	8.6	237	E55210	heti protein - Ana
6	85	8.6	1488	AH0615	cell division prot
7	84	8.5	292	C64699	proteinase IV (EC
8	83.5	8.4	542	A84229	sensory histidine
9	83	8.4	455	JC6051	trypsin-like prote
10	83	8.4	545	A87136	hypothetical prote
11	83	8.4	877	JN0772	glucan endo-1,3-be
12	82.5	8.3	446	S16308	translation elonga
13	82.5	8.3	522	D82195	phage replication
14	82	8.3	222	F71819	probable proteinas
15	82	8.3	788	A72330	phenylalanine-trNA
16	81	8.2	792	AB0704	pyruvate, water di
17	80.5	8.1	718	B81058	guanosine-3',5'-bi
18	80.5	8.1	725	E1819	guanosine-3',5'-bi
19	80.5	8.1	2117	T36180	GDA peptide synthe
20	80	8.1	536	F87297	feruloyl-CoA synth
21	80	8.1	1233	T40059	chromosome segrega
22	79	8.0	350	E87714	hypothetical prote
23	79	8.0	455	C91142	serine endoprotein
24	79	8.0	455	F85987	serine endoprotein
25	79	8.0	860	G82310	alanyl-tRNA synthe
26	79	8.0	1486	C64832	cell division prot
27	79	8.0	1486	G90754	kinesin-like cell
28	79	8.0	1486	E85618	hypothetical prote
29	79	8.0	1534	JH0228	cell division prot

30	78.5	7.9	322	2	F87549	type IV secretion
31	78.5	7.9	346	2	A48470	translation elonga
32	78.5	7.9	424	2	AC2975	oxidoreductase ord
33	78.5	7.9	424	2	H98307	probable oxidoredu
34	78.5	7.9	699	2	T18984	hypothetical prote
35	78.5	7.9	956	2	B71468	probable insulinase
36	78.5	7.9	968	2	T51933	kinesin motor prot
37	78.5	7.9	4684	2	A59404	plectin [imported]
38	78	7.9	325	2	A97482	hypothetical prote
39	78	7.9	325	2	A12699	conserved hypotet
40	78	7.9	335	2	C75619	histidinol-phospha
41	78	7.9	351	2	A83419	phage host specifi
42	78	7.9	366	2	G84849	probable actin [im
43	77.5	7.8	228	2	JN7467	ABC transporter, A
44	77.5	7.8	542	2	U00661	heat shock protein
45	77.5	7.8	542	2	S32106	groEL protein - La

ALIGNMENTS

RESULT 1

T17255
hypothetical protein DKFp586O1822.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T17255
R:Koehrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A:Reference number: Z18722
A:Accession: T17255
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1192 <KOE>
A:Cross-references: EMBL:AL117470
A:Experimental source: adult uterus; clone DKFzp586O1822
C:Genetics:
A:Note: DKFzp586O1822.1

Query Match	25.3%	Score	250;	DB 2;	Length	1192;
Best Local Similarity	44.5%	Pred. No.	1.7e-13;			
Matches	65;	Conservative	23;	Mismatches	48;	Indels 10; Gaps 4;
QY	54	DKLVS-FYLETYGAEILTANV--LRDMGLQEMA----	GOLQAATHQGSAAPIQAPPOS	106		
Db	1036	DQLFSEFYVGHLSGIRLQVKDKDETLVWEALVKPGLMPAT--	TLIPPARIAPVPSPL	1092		
QY	107	AAKPGHFIHQHRAALIAIARVTNVEWLLDALYKVLTDQYQAVRAPPTNP	SKMKLFSFT	166		
Db	1093	DAPQLLHFVDQYREQLIARVTSVEVVLDKLHGQVLSEQYERVLAENTRPSQ	MRKLFSL	1152		
QY	167	PAWNWTKDLLLQALRESQSYLVDEL	192			
Db	1153	QSWDRKCKDGLYQALKETHPHLIMEL	1178			

RESULT 2

T49723
hypothetical protein B23L21.360 [imported] - Neurospora crassa
C:Species: Neurospora crassa
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C:Accession: T49723
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25022
A:Accession: T49723
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-323 <SCH>
A:Cross-references: EMBL:AL356172; GSPDB:GN00116; NCSP:B23L21.360
A:Experimental source: BAC clone B23L21; strain OR74A
C:Genetics:
A:Gene: NCSP:B23L21.360

A;Map position: 6
A:Introns: 183/1: 229/2

Query Match	8.8%;	Score 87;	DB 2;	Length 323;
Best Local Similarity	20.7%;	Pred. No. 5.1;		
Matches	44;	Conservative 41;	Mismatches 80;	Indels 48; Gaps 9;
Qy	6	D A I L D A L E N L T A E L K K F K L K L L S V P R E G Y G R I P R G A L L S M D A L D L T D K L V S F Y L T Y G	65	
Db	119	D N H F D L L K V Q C G E D F E W F D E Q V K S L P - - - - - S P V T I E R F A T E N M A K H I V I R S G	168	
Qy	66	A E L T A N V L R D M G L O E M A G Q L O A A T H Q S G A P A G I O A P P O S A A K P G L H F I D Q H R A A L I A R	125	
Db	169	- - - - - E M S V R - M P G V V Y A - - H A G S T I L E V L R L Q V L V V Y P N P T L M D N H Q A E L - - -	213	
Qy	126	V T N V E W L D L A L Y G K V - - L T D E Q V A V R A E P T N P S K R K L F S F T P - - - - - A W N W	171	
Db	214	A Q E L E S T G D A V Y G R L G K L T A I A R S L E L V A Q G P L K L D L P Y S P P P P V P D S E R V T L F D W	273	
Qy	172	- - - - - T C - - - - - K O L L L Q A L R E S O S Y L V E D L E R S	195	
Db	274	M T L T C Y P D E L R K O O H L O D L A R N V E K S F O R E L E O A	306	

RESULT 3
 S23326
 gene ML2.2 protein precursor - Streptococcus pyogenes
 C;Species: Streptococcus pyogenes
 C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999
 C;Accession: S23326
 R;Bessen, D.E.; Fischetti, V.A.
 Infect. Immun. 60, 124-135, 1992
 A;Title: Nucleotide sequences of two adjacent M or M-like protein genes of group A streptococci
 A;Reference number: S23325; MUID:92104662; PMID:1370269
 A;Accession: S23326
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-372 <BES>
 A;Cross-references: EMBL:X61276; NID:g47369; PIDN:CAA43582.1; PID:g47371
 C;Superfamily: M5 protein

Query Match	8.7%; Score 86.5; DB 2; Length 372;
Best Local Similarity	26.8%; Pred. No. 6.7;
Matches	53; Conservative 30; Mismatches 78; Indels 37; Gaps 9;
QY	12 LENUABELKKFKLLKLVPLRGYGRIPGALLSMDAL-----DLTDKLVSFYLEYTGA 66 : : : : : : : : : : : :
Db	78 LEKINAEERKNKKLEANKELNENYYKLDG----IDALEKEREDLTKTTAKTTKNEIS 133 : : : : : : : : : : : :
QY	67 ELTANVL-RDMGLQEMAGQLQAATHQGSGAAPGIAPPQ--SAAKPGL-HFIDQHRAAL 122 : : : : : : : : : : : :
Db	134 EASRKGSLRDLEASRTAKKELEAKHOKLEAENKKLTGECNVSEASRKGLSNDLEASRAA- 192 : : : : : : : : : : : :
QY	123 IARVTNVVELLDALYGKVLTDQO-----YQAVRAEPTNPCKMKLFSPFTANNWTCOLL 177 : : : : : : : : : : : :
Db	193 -----KKELEAKYQKLETDHQDALEAKHKQLEADYQVSETSRKGLS-----RD--- 234 : : : : : : : : : : : :
QY	178 LQALRSQSYLVEDLERS 195 : : : : : :
Db	235 LEASREANKKVGTSELTOA 252 : : : : : :

RESULT 4
S75545
hypothetical protein slr0806 - *Synechocystis* sp. (strain PCC 6803)
C;Species: *Synechocystis* sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C;Accession: S75545
R;kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*

6.
A: Reference number: S74322; MUID: 97061201; PMID: 8905231

A:Accession: S75545
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-439 <XAN>
A:Cross-references: EMBL:D90911; GB:AB001339; NID:g1653083; PIDN:BAA18106.1; PID:g165319
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Superfamily: glycolate oxidase chain g1cd

[illegible]

```

RESULT 5
E55210
    C:Species: Anabaena sp. (strain PCC 7120)
    C:Date: 05-May-1995
    C:Accession: E55210
    R:Black, T.A.; Wolk, C.P.
    J: J. Bacteriol. 176, 2282-2292, 1994
    A:Title: Analysis of a Het- mutation in Anabaena sp. strain PCC 7120 im-
    A:Reference number: A55210; MUID:94209228; PMID:8157596
    A:Accession: E55210
    A:Status: preliminary
    A:Molecule type: DNA
    A:Residues: 1-237 <BLA>
    A:Cross-references: GB:L22883; NID:g441119; PIDN:AAA22003.1; PID:g441121
    C:Genetics:
    A:Gene: hetI
    A:Start codon: TTG

```

Query Match	8.6%; Score 85; DB 2; Length 237;
Best Local Similarity	23.7%; Pred. No. 5.2;
Matches	49; Conservative 32; Mismatches 72; Indels 54; Gaps 12
QY	14 NLTAELKKFKLLSVPLREGYGRIPRGALLSMDALDUTKLIV---SPVLEYTGAELTA 70
Db	12 NLT---LLSDVHLWRIPDQO-----PESQIQDLAATLSDELARANRYFFFEHRRRFTA 63
QY	71 --NVLRLMGLOEMAGQIOAATHOGSGAAPAGIOAPPOSAAKP-----GLHFDIQHR 119
Db	64 GRGILRSTI---LGGYL-----GVPEQGVFDFYESRGKPIGLDRFAESGLLFLNLSH 111
QY	120 AALITARVTN-----VEWL-----LDALYLGKVLTDQYQAVRAEPTNPCKMKLFSPT 166
Db	112 QNLALCAVNTROIQIDILEYLRPTSDLESIAKFFFLPREVELLRSUP-DEQKQKIFFRY- 169
QY	167 PAWNWTKOLLQALRESQSYLVLEDLE 193
Db	170 ----WTCKEAYLKATGDGIAGL-EEIE 191

RESULT 6
AH0615
cell division protein [imported] - Salmonella enterica subsp. enterica serovar Typhi (str. C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence revision 09-Nov-2001 #text change 18-Nov-2002

[illegible]

QY 11 ALNLTAEELKFKLLKLVPLREGYGRIPRG-ALLSMDALDLTKLVSF-----YLET 63
 DB 217 ALLNLNGE-----LIGINTAILAPGGSGVIGFAIPSNWARTLAQQLIDFGIKRGLGI 271
 QY 64 YGAELTANVLDMGLQEMAGOLQAATHQSGGAAPAGIQIA-----102
 DB 272 KGTMSADIATAFNLDVQRFVSEVLPGSGSAKAGVAGDIITSLNGKPLNSFAELRSR 331
 QY 103 -----PPQSAKPGGLHFIDQHRAALIAVIT-----NVELLDALYKVLTDQY--- 146
 DB 332 IATTEPGTKVKGLL-----RNGKPLEVEVLTDTSTSSASAEMITPALEGATILSDGQLKDG 388
 QY 147 -----QAVRAEP-----TNPSKMKLFSTTPAANNWTKDILLQ 179
 DB 389 GGIKIDVVGKSPAAQAGLQKDDVVIIVGNRRVNSIAEMRKVLAAPKA-----IALQ 442
 QY 180 ALRESQS-YLV 189
 DB 443 IVRGNESIYLL 453
 RESULT 10
 A:Title: hypothetical protein ML1816 [imported] - Mycobacterium leprae
 C:Species: Mycobacterium leprae
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
 C:Accession: A87136
 R.; Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho
 R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
 eam, M.A.; Rutherford, K.M.
 Nature 409, 1007-1011, 2001
 A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S
 A:Title: Massive gene decay in the leprosy bacillus.
 A:Reference number: A86909; PMID:21128732; PMID:11234002
 A:Accession: A87136
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-545 <STO>
 A:Cross-references: GB:AL450380; NID:gl3093526; PIDN:CAC30769.1; GSPDB:GN00147
 A:Gene: ML1816
 C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
 Query Match 8.4%; Score 83; DB 2; Length 545;
 Best Local Similarity 24.4%; Pred. No. 22;
 Matches 41; Conservative 15; Mismatches 50; Indels 62; Gaps 7;
 QY 51 DLTKLVSPY-----LETYGAELTA-NVLDMGLQ-----EM 81
 DB 113 DARDRAIRYQLEERFLALGGYSAEAGRICASGLPERVLVQQLCTLSGGQRRVEL 172
 QY 82 AGQLQAATHQSGGAAPAG-----IQAPPSAAKPGGLHF-----IDQHRAALIA 124
 DB 173 ARILPAASRAGTAGSGGTTLLDDEPTNHLDAADSLGWLDFLRSHRTGGLWISHNVLLIA 232
 QY 125 RVTNVELLDALYKGV-----LTDQYQAVRAEPTNSK 158
 DB 233 AVNVRVWFDAVLGKVDVYVMGMGYKYLDSSRATDEQRR--RRERVNAER 278
 RESULT 11
 JN0772
 glucan endo-1,3-beta-D-glucosidase (EC 3.2.1.39) precursor - Bacillus circulans (strain
 N:Alternate names: beta-1,3-glucanase
 C:Species: Bacillus circulans
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 21-Jul-2000
 C:Accession: JN0772; P0613
 R.; Yamamoto, M.; Aono, R.; Horikoshi, K.
 Biosci. Biotechnol. Biochem. 57, 1518-1525, 1993
 A:Title: Structure of the 87-kDa beta-1,3-glucanase gene of Bacillus circulans. IAMI165 a
 A:Reference number: JN0772; MUID:94033858; PMID:7764221
 A:Accession: JN0772

A:Molecule type: DNA
 A:Residues: 1-877 <YAM>
 A:Cross-references: GB:DI7519; NID:9601877; PIDN:BA04469.1; PID:9601878
 A:Accession: P0613
 A:Molecule type: protein
 A:Residues: 39-58 <YA2>
 C:Genetics:
 A:Gene: bglH
 C:Keywords: glycosidase; hydrolase; polysaccharide degradation
 F:1-38/Domain: signal sequence #status predicted <SIG>
 F:39-877/Product: glucan endo-1,3-beta-D-glucosidase #status predicted <MAT>
 Query Match 8.4%; Score 83; DB 2; Length 877;
 Best Local Similarity 22.3%; Pred. No. 39;
 Matches 40; Conservative 26; Mismatches 73; Indels 40; Gaps 8;
 QY 10 DALENLTAEELKFKLLKLVPLR-----EGYGRIPRGALLSMDALDLTKLVSYLE-- 62
 DB 471 DRAQNVFVQD-GKLNKALNEPKSPQDPSPRYAQYSSGKINTKDFSLKYGRVDFRAKLIP 529
 QY 63 -----TYGAELTANVLDMGLQEMAGOLQAATHQSGGAAPAGIQAPPS 106
 DB 530 TNGIWPALWMLPQDNVYGTWASSG---EIDVMEAKGRLPST---SGAVHFGGQWPTNR 583
 QY 107 AAKPGLHF-----IDQHRAALIAVITNVELLDLA-LYGVKVLTDQYQAVRAEPTNS 157
 DB 584 YLSGEVHFPEGQTFFANDYHYVSVVWEEDNIKWYVDGKFFKVTTRDQWYSA--AAPNPN 640
 RESULT 12
 S16308
 translation elongation factor eEF-1 alpha chain - Stylonychia lemnae
 C:Species: Stylonychia lemnae
 C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 02-Feb-2001
 C:Accession: S16308
 R.; Bierbaum, P.; Doenhoff, T.; Klein, A.
 Mol. Microbiol. 5, 1567-1575, 1991
 A:Title: Macronuclear and micronuclear configurations of a gene encoding the protein synt
 A:Reference number: S16308; MUID:92157884; PMID:1840642
 A:Accession: S16308
 A:Molecule type: DNA
 A:Residues: 1-446 <BIE>
 A:Cross-references: EMBL:X57926; NID:gl10139; PIDN:CAA41001.1; PID:gl10140
 A:Note: the authors translated the codon GAT for residue 307 as Gly
 C:Genetics:
 A:Gene: efaa
 A:Genetic code: SGC5
 C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolog
 C:Keywords: GTP binding; nucleotide binding; p-loop; protein biosynthesis
 F:8-156/Domain: translation elongation factor Tu homology <ETU>
 F:14-21/Region: nucleotide-binding motif A (p-loop)
 F:153-156/Region: GTP-binding NKXD motif
 Query Match 8.3%; Score 82.5; DB 2; Length 446;
 Best Local Similarity 21.4%; Pred. No. 19;
 Matches 44; Conservative 40; Mismatches 81; Indels 41; Gaps 8;
 QY 8 ILDALENLTAEELKFKLLKLVPLREGY-----GRIPRGAL-----LSMDALDL 52
 DB 216 LIDALDAL--DQKPKDKPLRLPLQDYVKIGIGITVPVGRVETGLLKGMLVTFAPMNI 273
 QY 53 TDKLVSYLETYGAELTA-----NV---LRDMGLQEMAGOLQAATHQSGGA-----AP 97
 DB 274 TTECKS--VEMHESLTAEPGDNDVGVTKVLSVDLGRGVYASDKNDPAKDTTNFLAQ 331
 QY 98 AGIQAPPSAAKPGGLHFIDQHRAALIAVITNVELLDALYKVLTD-----QYQAVR 150
 DB 332 VIVLHPGQIQGYAPVLDCHTAHACKDFDEIESKVDRRSGKVLSEEPKFKSGEAALVR 391
 QY 151 APTNPSKMKLFSTTPAANNWTKDIL 176
 DB 392 MYVQPKMCMVEAFNQYPPILGRFAVRDM 417

```

RESULT 13
D82195
phage replication protein Cri VC1469 VC1475 [imported] - Vibrio cholerae (strain N16961
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: D82195; A82196
R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F.
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: D82195
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-522 <HE1>
A:Cross-references: GB:AE004225; GB:AE003852; NID:99555968; PIDN:AAF94626.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
A:Accession: A82196
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-522 <HE2>
A:Cross-references: GB:AE004225; GB:AE003852; NID:99555968; PIDN:AAF94631.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics: <GEN1>
A:Gene: VC1469
A:Map position: 1
C:Genetics: <GEN2>
A:Gene: VC1475
A:Map position: 1

Query Match 8.3%; Score 82.5; DB 2; Length 522;
Best Local Similarity 24.4%; Pred. No. 23;
Matches 38; Conservative 20; Mismatches 53; Indels 45; Gaps 7;

QY 10 DALENITAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLTKLVSFYLETYCAEL 68
DB 300 DLLESKLGEELELYNDSKIKQLLRDAYSTWTPKGNISYAKA-----DRLFRFYM----- 348

QY 69 TANVLDMGLQEMAGLOAQAT-HQGS-----GAAPAGIQAPQSAKPKLHFIDQHRAA 121
DB 349 ---TLCDRGYQELKAHSSKATLRNMRDLMAIGFSKADLQNLSEGRMP----- 394

QY 122 LIAATVNVEMLLDALYKGLVTDQYQVRAEPTNPS 157
DB 395 -LAQVLNFN-----DNQRPANVVEPVST 418

RESULT 14
F71819
probable proteinase IV - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: Strain J99
C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 29-Sep-1999
C:Accession: F71819
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: F71819
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-292 <ARN>
A:Cross-references: GB:AE001556; GB:AE001439; NID:g4155938; PIDN:AAD06917.1; PID:g415595
A:Experimental source: strain J99
C:Genetics:
A:Gene: sppA
C:Superfamily: proteinase IV MTH806

Query Match 8.3%; Score 82; DB 2; Length 292;
Best Local Similarity 23.0%; Pred. No. 12;

```

```

Matches 40; Conservative 32; Mismatches 60; Indels 42; Gaps 7;

QY 42 GALLSMDALDL-TDKL-----VSFYLETYGAELTANV-----LRDMGLQEMAGLOQA 87
DB 57 GAIFSTEDFDKEVDKILKTPSIKGVLLIDSPGASVASVELSEKIALDL-KQKMPVLA 115
QY 88 ATHQSGAAPAGIQAPQSAKPKLHFIDQHRAAIARV-----TNVENLLDALYKGLV 141
DB 116 RGVWASGSGYAGMQASVEYASK-----ASLIGSIGVIFSSANVENLANKV----- 160
QY 142 TDEQYQVRAEPTNPSKMKLFSTPAWNWTKDQLLQALRESQSVLYVEDLERS 195
DB 161 -----GVATQGVHAGEYKEIGTFTFRANKPNEKEFLQNLVNEQYQMFVDDVAKA 208

```

RESULT 15

A72330

```

phenylalanine-tRNA ligase (EC 6.1.1.20) beta chain - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 03-Jun-2002
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Hatt, D.H.; Hickey,
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: A72330
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-788 <ARN>
A:Cross-references: GB:AE001749; GB:AE000512; NID:g4981346; PIDN:AAD35904.1; PID:g498135;
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM0822

```

```

C:Superfamily: phenylalanine-tRNA ligase beta chain
C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 8.3%; Score 82; DB 2; Length 788;
Best Local Similarity 24.4%; Pred. No. 42;
Matches 29; Conservative 19; Mismatches 57; Indels 14; Gaps 2;

QY 6 DAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLTKLVSFYLETYG 65
DB 87 DVVILALEGATLN-----NGLKIEPRFKGVISEGMLCSLEELGLEKSDRVVRFDP 139
QY 66 AEILTANVLDMGLQEMAGLOAQATHQ-----GSAAPAGIQAPQSAKPKLHFIDQ 117
DB 140 VELGVNWEYGLNERVLDEITPNRDCLSITIGVARELSALSGRPLNKQPDVSVFDE 198

```

Search completed: January 29, 2004, 13:49:47
Job time : 6.80296 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:40:41 ; Search time 3.00185 Seconds
(without alignments)
3054.855 Million cell updates/sec

Title: US-09-996-617-8
Perfect score: 990
Sequence: 1 MGRARDAILDALENLTABEL.....LLQLALRESQSYLVEDLERS 195

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	990	100.0	195	1 ASC_HUMAN	Q9ulz3 homo sapien
2	684	69.1	193	1 ASC_MOUSE	Q9epb4 mus musculu
3	250	25.3	1473	1 NAL1_HUMAN	Q9c000 homo sapien
4	213	21.5	203	1 ASC_BRARE	Q9i9n6 brachydanio
5	138	13.9	854	1 PYA5_RAT	Q63035 rattus norv
6	132	13.3	843	1 PYA5_MOUSE	Q91ws2 mus musculu
7	125.5	12.7	892	1 PYA5_HUMAN	P59044 homo sapien
8	122.5	12.4	781	1 MEV3_HUMAN	O15553 homo sapien
9	99.5	10.1	980	1 PYA3_HUMAN	Q8wx94 homo sapien
10	99	10.0	431	1 CAR8_HUMAN	Q9Y2q2 homo sapien
11	91.5	9.2	143	1 SILE_SALTY	Q9Z4n3 salmonella
12	89	9.0	338	1 RUVE_THETN	Q8ran2 thermoanaer
13	86.5	8.7	372	1 M22_STRPY	P50469 streptococc
14	86.5	8.7	835	1 SYFB_CORGL	Q8nqn6 corynebacte
15	86	8.7	1033	1 C1S1_MOUSE	Q8r4b8 mus musculu
16	85	8.6	237	1 HET1_ANASP	P37695 anabaena sp
17	84.5	8.5	540	1 RIK2_HUMAN	O43353 h receptor-
18	83	8.4	278	1 ACT_SCHDU	P65314 scherffelia
19	83	8.4	455	1 DEGO_ECOLI	P39099 escherichia
20	82.5	8.3	446	1 SYFB_STYLE	P25166 stylonychia
21	82	8.3	788	1 SYFB_THEMEA	Q9wz89 thermotoga
22	80	8.1	539	1 RIK2_MOUSE	P58801 mus musculu
23	80	8.1	1034	1 C1S1_HUMAN	Q96p20 homo sapien
24	80	8.1	1233	1 SMCI_SCHPO	Q94583 schizosacch
25	79	8.0	860	1 SYA_VIECH	P56648 vibrio chol
26	79	8.0	1486	1 MUKB_ECOLI	Q22523 escherichia
27	78.5	7.9	346	1 EPIA_EIMBO	Q107051 eimeria bov
28	78.5	7.9	4684	1 PLEI_HUMAN	Q15149 homo sapien
29	78	7.9	366	1 ACT9_ARATH	P93738 arabidopsis
30	76.5	7.7	213	1 P914_THEMEA	Q9x016 thermotoga
31	76.5	7.7	324	1 PARB_TREPA	O83295 treponema p
32	76.5	7.7	953	1 CAR4_HUMAN	Q9Y239 homo sapien
33	76	7.7	739	1 PURL_LISMO	Q8Y6c1 listeria mo

34	75.5	7.6	444	1	TIG-AZOBR	Q9x6w7 azospirillu
35	75.5	7.6	542	1	CH60_LACLA	P37282 lactococcus
36	75.5	7.6	542	1	CH60_LACLC	Q9asp7 lactococcus
37	75.5	7.6	546	1	CH60_LACAC	Q93g07 lactobacill
38	75.5	7.6	611	1	BIR_CHICK	Q90660 gallus gall
39	75.5	7.6	1698	1	Y076_HUMAN	Q14999 homo sapien
40	75	7.6	545	1	CH60_PARDE	Q92462 paracoccus
41	75	7.6	791	1	PFSA_ECOLI	P23538 escherichia
42	74.5	7.5	401	1	ISPE_LYCES	P93841 lycopersico
43	74.5	7.5	862	1	LOXA_PHAVU	P27480 phaseolus v
44	74.5	7.5	957	1	SECA_MYCSM	P71533 mycobacteri
45	74.5	7.5	1062	1	PIA7_HUMAN	P59046 homo sapien

RESULT 1

ASC_HUMAN ID - ASC_HUMAN STANDARD; PRT; 195 AA.

AC Q9ULZ3; Q96D12; Q9BSZ5; Q9HBD0; Q9NXJ8; DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Apoptosis-associated speck-like protein containing a CARD (hASC) (PYCARD) (Target of methylation-induced silencing 1) (Caspase DE recruitment domain protein 5).

GN ASC OR TMS1 OR CARD5.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]_TaxID=9606;

RP SEQUENCE FROM N.A. (ISOFORM 1).

RC TISSUE=Leukemia;

RX MEDLINE=20036508; PubMed=10567338; RA Masumoto J., Taniguchi S., Ayukawa K., Sarvotham H., Kishino T., RA Niihara N., Hidaka E., Katsuyama T., Higuchi T., Sagara J.; RT "ASC, a novel 22-kDa protein, aggregates during apoptosis of human RT promyelocytic leukemia HL-60 cells," J. Biol. Chem. 274:33835-33838(1999).

RL [2]

RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).

RP TISSUE=Fibroblast;

RX MEDLINE=20552139; PubMed=11103776; RA Conway K.E., McConnell B.B., Bowring C.E., Donald C.D., Warren S.T., RA Vertino P.M.; RT "TMS1, a novel proapoptotic caspase recruitment domain protein, is a RT target of methylation-induced gene silencing in human breast RT cancers," Cancer Res. 60:6236-6242(2000).

RL [3]

RN SEQUENCE FROM N.A. (ISOFORM 1).

RP Martinon F., Hofmann K., Tschopp J.; RT "Pycard a PYD and CARD containing molecule," RT Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

RL [4]

RN SEQUENCE FROM N.A. (ISOFORM 1).

RP Bertin J.; RT "CARD5 protein is a CARD/PYRIN family member that is involved in RT apoptosis signal transduction," RT Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

RL [5]

RN SEQUENCE FROM N.A. (ISOFORM 1).

RP TISSUE=Colon mucosa;

RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T., RA Suzuki Y., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., RA Nakamura Y., Isoigai T., Sugano S.; RT "NEO human cDNA sequencing project," RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

RL [6]

RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).

RP TISSUE=Lymph, and Pancreas;

ALIGNMENTS

RA MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Tohiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J.J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [?]
RP FUNCTION, AND SUBCELLULAR LOCATION.
RP MEDLINE=20552140; PubMed=11103777;
RX McConnell B.B., Vertino P.M.;
RT "Activation of a caspase-9-mediated apoptotic pathway by subcellular
RT redistribution of the novel caspase recruitment domain protein TMS1";
RL Cancer Res. 60:6243-6247(2000).
CC -!- FUNCTION: Promotes caspase-mediated apoptosis. This proapoptotic
CC activity is mediated predominantly through the activation of
CC caspase 9.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic. Upstream of caspase activation,
CC a redistribution from the cytoplasm to the aggregates occurs.
CC These appeared as hollow, perinuclear spherical, ball-like
CC structures.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q9ULZ3-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9ULZ3-2; Sequence=VSP_004119;
CC Name=3;
CC IsoId=Q9ULZ3-3; Sequence=VSP_004118;
CC Note=Inferred from the cDNA sequence of Ref. 5;
CC -!- TISSUE SPECIFICITY: Widely expressed at low levels. Detected in
CC peripheral blood leukocytes, lung, small intestine, spleen,
CC thymus, colon and at lower levels in placenta, liver and kidney.
CC Very low expression in skeletal muscle, heart and brain. Detected
CC in the leukemia cell lines HL-60 and U937, but not in Jurkat T-
CC cell lymphoma and Daudi Burkitt's lymphoma. Detected in the
CC melanoma cell line WM35, but not in WM793. Not detected in HeLa
CC cervical carcinoma cells and Molt 4 lymphocytic leukemia cells.
CC -!- MISCELLANEOUS: In breast tumorigenesis, methylation-mediated
CC silencing may affect genes and proteins that act as positive
CC mediators of cell death.
CC -!- SIMILARITY: BELONGS TO THE CARD-CONTAINING ADAPTER PROTEIN FAMILY.
CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 1 CARD domain.
CC -!- CAUTION: Ref.5 sequence differs from that shown due to a
CC frameshift in position 4.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB023416; BAA87339.2; -;
DR EMBL; AF184072; AAG01187.1; -;
DR EMBL; AF184073; AAG01188.1; -;
DR EMBL; AF255794; AAF99665.1; -;

DR EMBL; AF310103; AAG30286.1; -;
DR EMBL; AF384665; AAK63850.1; -;
DR EMBL; AK000211; BAA91012.1; ALT_FRAME.
DR EMBL; BC004470; AAH04470.1; -;
DR EMBL; BC013569; AAH13569.1; ALT_INIT.
DR MIM; 606838; -;
DR GO; GO:0006917; P:induction of apoptosis; TAS.
DR InterPro; IPR001315; CARD.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF02758; PAAD_DAPIN; 1.
DR PROSITE; PS0209; CARD; 1.
DR PROSITE; PS0824; DAPIN; 1.
KW Apoptosis; Anti-oncogene; Alternative splicing.
FT DOMAIN 1 91
FT DAPIN. CARD.
FT VARSPLIC 107 195
FT Missing (in isoform 3).
FT VARSPLIC 93 111
FT Missing (in isoform 2).
FT /FTID=VSP_004119.
SQ SEQUENCE 195 AA; 21627 MW; 455987286586F46A CRC64;

Query Match 100.0%; Score 990; DB 1; Length 195;
Best Local Similarity 100.0%; Pred. No. 4.8e-79;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRARDAILDALENTAEELKFKLLSVLPUREGYGRIPRGALLSMDALDLDTKLVSPY 60
DB 1 MGRARDAILDALENTAEELKFKLLSVLPUREGYGRIPRGALLSMDALDLDTKLVSPY 60
QY 61 LETYGAELTANYLRDMGLQEMAGQLQAATHQSGGAAPAGIQAPPOSAAKPGUHFIDQHRA 120
DB 61 LETYGAELTANYLRDMGLQEMAGQLQAATHQSGGAAPAGIQAPPOSAAKPGUHFIDQHRA 120
QY 121 ALIARTVNVWLLDALYKGLTDEQYQAVRAEPTNPVKRLFSFTPAWNWTKDILLQA 180
DB 121 ALIARTVNVWLLDALYKGLTDEQYQAVRAEPTNPVKRLFSFTPAWNWTKDILLQA 180
QY 181 LRESQSYLVEDLERS 195
DB 181 LRESQSYLVEDLERS 195

RESULT 2
ASC_MOUSE STANDARD; PRT; 193 AA.
ID ASC_MOUSE STANDARD; PRT; 193 AA.
AC Q9EFB4; Q9D2W9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Apoptosis-associated speck-like protein containing a CARD (mASC)
DE (PYCARD).
GN ASC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=Thymus;
RX MEDLINE=20580347; PubMed=11139337;
RA Masumoto J., Taniguchi S., Nakayama K., Ayukawa K., Sagata J.;
RT "Murine ortholog of ASC, a CARD-containing protein, self-associates
RT and exhibits restricted distribution in developing mouse embryos.";
RL Exp. Cell Res. 262:128-133(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Breast tumor;
RA Martinon F., Hofmann K., Tschopp J.;
RT "pycard a PYD and CARD containing molecule";
RL Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Pancreas, and Tongue;

RX MEDLINE=21085660; PubMed=11217951;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Akawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
 RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bull D., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Tovo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690 (2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL human and mouse cDNA sequences.";
 CC -!- FUNCTION: Promotes caspase-mediated apoptosis. This proapoptotic
 activity is mediated predominantly through the activation of
 caspase 9 (by similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic. Upstream of caspase activation,
 a redistribution from the cytoplasm to the aggregates occurs.
 CC These appeared as hollow, perinuclear spherical, ball-like
 structures (by similarity).
 CC -!- TISSUE SPECIFICITY: Expressed in small intestine, colon, thymus,
 spleen, brain, heart, skeletal muscle, kidney, lung and liver.
 CC -!- DEVELOPMENTAL STAGE: Strongly expressed at E9.5 day in the
 telencephalon, thalamic areas of the diencephalon, heart and
 liver.
 CC -!- SIMILARITY: BELONGS TO THE CARD-CONTAINING ADAPTER PROTEIN FAMILY.
 CC -!- SIMILARITY: Contains 1 DAPIN domain.
 CC -!- SIMILARITY: Contains 1 CARD domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB032249; BAB16609.1; -;
 DR EMBL; AF310104; BAG30287.1; -;
 DR EMBL; AK009852; BAB26543.1; -;
 DR EMBL; AK007742; BAB25229.1; -;
 DR EMBL; AK018682; BAB31341.1; -;

DR EMBL; BC008252; AAH08252.1; -;
 DR MGD; MGI:1931465; Asc.
 DR GO; GO:0005829; C:cytosol; IDA.
 DR InterPro; IPR001315; CARD.
 DR InterPro; IPR004020; PAAD-DAPIN_dom.
 DR Pfam; PF02758; PAAD-DAPIN; 1.
 DR PROSITE; PS50209; CARD; 1.
 DR PROSITE; PS50824; DAPIN; 1.
 KW Apoptosis; Anti-oncogene.
 FT DOMAIN 1 91 DAPIN.
 FT DOMAIN 105 193 CARD.
 FT CONFLICT 159 159 K -> E (IN REF. 3).
 FT CONFLICT 159 159 K -> E (IN REF. 3).
 SQ SEQUENCE 193 AA; 21458 MW; 244EA40194870B31 CRC64;
 Query Match 69.1%; Score 684; DB 1; Length 193;
 Best Local Similarity 71.8%; Pred. No. 1.8e-52;
 Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;
 QY 1 MGRARDAILDALENUTABELKFKLLSVLPREGYGRIPRGALLSMDALDLDTKLVSPY 60
 DB 1 MGRARDAILDALENLSGDELKFKLLTVQREGYGRIPRGALLQMDAIDLTKLVSY 60
 QY 61 LETYGAELTANVLRDMGLQEMAGOLQAATHQSGGAAPAGIQAPPOSAARPGHLFIDQRA 120
 DB 61 LESYGLLELTVLRDMGLQELAEQLQ-TTKESGAVAAASVPAQSTARTG-HFVDQHRQ 118
 QY 121 ALIARVTVNVEMLDALYGVLTDEQYQAVRABPTNPKNRKLFSFTPAWNWTKOLLQA 180
 DB 119 ALIARVTEVDGVLDAHGSVLTGQYQAVRAETTSQDKMRKLFSPVPSNNLTCKDSLQA 178
 QY 181 LRESOSYLVEDLERS 195
 DB 179 LKEIHPYLVMDLEQS 193
 RESULT 3
 NALI HUMAN
 ID NALI HUMAN STANDARD; PRT; 1473 AA.
 AC Q9C000; Q9BZ28; Q9HAV8; Q9UFT4; Q9Y2E0;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE NACTH-, LRR- and PYD-containing protein 2 (Death effector filament-
 forming ced-4-like apoptosis protein) (Nucleotide-binding domain and
 caspase recruitment domain) (Caspase recruitment domain protein 7).
 GN NALP1 OR DEFCAP OR NAC OR CARD7 OR KIAA0926.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=21169419; PubMed=11270363;
 RA Bertin J., DiStefano P.S.;
 RT "The PYRIN domain: a novel motif found in apoptosis and inflammation
 proteins.";
 RL Cell Death Differ. 7:1273-1274 (2000).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=21148093; PubMed=11250163;
 RA Martinon F., Hofmann K., Tschopp J.;
 RT "The pyrin domain: a possible member of the death domain-fold family
 implicated in apoptosis and inflammation.";
 RL Curr. Biol. 11:R118-R120 (2001).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RX TISSUE=Mythroleukemia;
 RC MEDLINE=21153743; PubMed=11076957;
 RA Hlaing T., Guo R.-F., Dilley K.A., Loussia J.M., Morrish T.A.,
 RA Shi M.M., Vincenz C., Ward P.A.;
 RT "Molecular cloning and characterization of DEFCAP-L and -S, two
 isoforms of a novel member of the mammalian Ced-4 family of apoptosis
 proteins.";

J. Biol. Chem. 276:9230-9238(2001).

[4] SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3 AND 4), AND PROTEIN INTERACTION.

TISSUE=T-cell;

RA MEDLINE=21153744; PubMed=11113115;

RA Chu Z.-L., Pio F., Xie Z., Welsh K., Krajewski M., Krajewski S.,

RA Godzik A., Reed J.C.;

RT "A novel enhancer of the Apaf1 apoptosome involved in cytochrome

RT c-dependent caspase activation and apoptosis.";

RL J. Biol. Chem. 276:9239-9245(2001).

[5] SEQUENCE FROM N.A. (ISOFORM 2).

TISSUE=Brain;

RA MEDLINE=99246063; PubMed=10231032;

RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M.,

RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;

RT "Prediction of the coding sequences of unidentified human genes. XIII.

RT The complete sequences of 100 new cDNA clones from brain which code

RT for large proteins in vitro.";

RL DNA Res. 6:63-70(1999).

[6] SEQUENCE OF 282-1473 FROM N.A. (ISOFORM 1).

TISSUE=Uterus;

RA Koehrer K., Beyer A., Mewes H.-W., Gassenhuber J., Wiemann S.;

RL Submitted [SEP-1999] to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: Able to form cytoplasmic structures termed death

CC effector filaments. Enhances APAF1 and cytochrome c-dependent

CC activation of pro-caspase-9 and consecutive apoptosis. Seems to

CC bind ATP.

CC -!- SUBUNIT: Interacts strongly with caspase 2, weakly with caspase 9

CC and with APAF1 in a cytochrome c-inducible way leading to the

CC formation of an apoptosome. This interaction may be ATP-dependent.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=4;

CC Name=1; Synonym=NAC beta, DEFCAP-L;

CC IsoId=O9C000-1; Sequence=Displayed;

CC Name=2; Synonym=NAC alpha, DEFCAP-S;

CC IsoId=O9C000-2; Sequence=VSP_004327;

CC Name=3; Synonym=NAC gamma;

CC IsoId=O9C000-3; Sequence=VSP_004326, VSP_004327;

CC Name=4; Synonym=NAC delta;

CC IsoId=O9C000-4; Sequence=VSP_004326;

CC -!- TISSUE SPECIFICITY: Widely expressed. Isoforms 1 and 2 are

CC expressed in peripheral blood leukocytes, chronic myelogenous

CC leukemia cell line K-562, followed by thymus, spleen and heart.

CC Also detected in lung, placenta, small intestine, colon, kidney,

CC liver and muscle.

CC -!- SIMILARITY: Contains 1 DAPIN domain.

CC -!- SIMILARITY: Contains 1 NACHT domain.

CC -!- SIMILARITY: Contains 1 CARD domain.

CC -!- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AF298548; AAG15254.1; -

DR EMBL: AF310105; AAG30288.1; -

DR EMBL: AF229059; AAK00748.1; -

DR EMBL: AF229060; AAK00749.1; -

DR EMBL: AF229061; AAK00750.1; -

DR EMBL: AF229062; AAK00751.1; -

DR EMBL: AB023143; BA076770.1; -

DR EMBL: AL114701; CAB55945.1; -

DR PIR: T17255; T17255.

DR HSP: P13489; I4Y4.

DR MIM: 606636; -

GO; GO:0005622; C:intracellular; IC.

GO; GO:0016506; F:apoptosis activator activity; NAS.

GO; GO:0008656; F:caspase activator activity; NAS.

GO; GO:0019899; F:enzyme binding activity; IPI.

GO; GO:0006915; F:caspase activation; NAS.

GO; GO:0006917; F:induction of apoptosis; NAS.

InterPro; IPR001315; CARD.

InterPro; IPR000767; Disease_resist.

InterPro; IPR001611; LRR.

InterPro; IPR007091; LRR_RNinh.

InterPro; IPR007111; NACHT_NTPase.

InterPro; IPR004020; PAAD_DAPIN_dom.

Pfam; PF00560; LRR; 2.

Pfam; PF02758; PAAD_DAPIN; 1.

PRINTS; PR00364; DISEASERSIST.

PROSITE; PS02029; CARD; 1.

PROSITE; PS0824; DAPIN; 1.

PROSITE; PS0837; NACHT; 1.

Apoptosis; ATP-binding; Leucine-rich repeat; Repeat;

Alternative splicing.

DOMAIN 1 92

DOMAIN 328 637

DOMAIN 704 725

REPEAT 807 830

REPEAT 864 887

REPEAT 921 944

REPEAT 950 973

REPEAT 1139 1215

REPEAT 1216 1236

DOMAIN 1374 1463

NP_BIND 334 341

VARSPLIC 958 987

VARSPLIC 1262 1305

MUTAGEN 340 340

MUTAGEN 340 340

CONFLICT 155 155

CONFLICT 246 246

CONFLICT 782 782

CONFLICT 878 878

CONFLICT 995 995

CONFLICT 1119 1119

CONFLICT 1184 1184

CONFLICT 1241 1241

CONFLICT 1366 1366

SEQUENCE 1473 AA; 165865 MW; 438F0DC45C2562D CRC64;

Query Match 25.3%; Score 250; DB 1; Length 1473;

Best Local Similarity 44.5%; Pred. No. 9.6e-14;

Matches 65; Conservative 23; Mismatches 48; Indels 10; Gaps 4;

QY 54 DKLVSYFETVYGAELTANV--LRDMGLQEMA-----GQLQATHQSGNAPAGIQAPPOS 106.

Db 1317 DQLFSEFYVGHLSGIRLQVKDKKDETLVWEALVKGDLMPAT---TLIPPARIAVPSPL 1373

QY 107 AAKPGHFDHRAALIAVTNVEMLLDALVGKVLTDQYQAVRAEPTNPSKMKLFSFT 166

Db 1374 DAPQLHFDVQREGLIARVTSVEVLDKLGQVLSQEQYERVLAEINTRPSQMKLFLSL 1433

QY 167 PAWNWTCKDLLLQALRESQSYLVEDL 192

Db 1434 QSMRDKCKDGLYQALKETHPHLIMEL 1459

RESULT 4

ASC BRARE

ID ASC BRARE STANDARD; PRT; 203 AA.

AC Q913N6

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Apoptosis-associated speck-like protein containing a CARD (PYCARD).

GN ASC OR ASC1.

Matches 32; Conservative 17; Mismatches 25; Indels 4; Gaps 2;
QY 4 ARDAILDALENITAEELKKFKLLSVPLREGYGR-IPRGALLSMDALDLDTKLVSYFYLE 62
DB 15 ARELLAALQDLSQEQQLKFRHKLRDAPLD---GRSIPWGRLESHSDAVALDVKLIEFYAP 71
QY 63 TYGAELTANVLRLDMGLQE 80
DB 72 EPAVDVTRKLLKCADIRD 89
RESULT 6
PYAS MOUSE
ID PYAS5 MOUSE STANDARD; PRT; 843 AA.
AC Q91W52: Q8K0L4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE PYRIN-containing APAF1-like protein 5-like.
GN PYPAF5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP CONCEPTUAL TRANSLATION OF 1-287.
RA Hinz U.;
RL Unpublished observations (FEB-2003).
RN [2]
RP SEQUENCE OF 288-843 FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP IDENTIFICATION OF MAMMALIAN ORTHOLOGS OF PYPAF5.
RA Albrecht M., Domingues F.S., Schreiber S., Lengauer T.;
RT "Identification of mammalian orthologs associates PYPAF5 with distinct
functional roles";
RL FEBS Lett. 538:173-177(2003).
CC -!- FUNCTION: May mediate activation of CASP1 via ASC and promote
activation of NF-kappa-B (By similarity).
CC -!- SUBUNIT: Binds to ASC with its DAPIN domain (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- SIMILARITY: Contains 3 leucine-rich (LRR) repeats.
CC -!- CAUTION: The N-terminus was extended using ESTs and genomic
sequences, in analogy to ortholog sequences.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial

entites requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; BC013519; AAH13519.1; -
CC EMBL; BC011139; AAH11139.1; ALT_INIT.
DR MGD; MGI:2141990; Pypaf5.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh.
DR Pfam; PF00560; LRR; 1.
DR PROSITE; PS50824; DAPIN; 1.
DR PROSITE; PS50837; NACHT; 1.
KW ATP-binding; Leucine-rich repeat; Repeat.
FT DOMAIN 11 102 DAPIN.
FT DOMAIN 168 485 NACHT.
FT DOMAIN 569 585 ASP-GLU-RICH.
FT DOMAIN 654 661 POLY-LYS.
FT REPEAT 434 459 LRR 1.
FT REPEAT 609 632 LRR 2.
FT REPEAT 811 834 LRR 3.
FT NP BIND 174 181 ATP (POTENTIAL).
SQ SEQUENCE 843 AA; 94592 MW; 35FB7A766A47DB51 CRC64;
Query Match 13.3%; Score 132; DB 1; Length 843;
Best Local Similarity 41.0%; Pred. No. 0.00089;
Matches 32; Conservative 14; Mismatches 28; Indels 4; Gaps 2;
QY 4 ARDAILDALENITAEELKKFKLLSVPLREGYGR-IPRGALLSMDALDLDTKLVSYFYLE 62
DB 15 ARELLAALQDLSQEQQLKFRHKLRDAPLD---GRSIPWGRLESHSDAVALDVKLIEFYAP 71
QY 63 TYGAELTANVLRLDMGLQE 80
DB 72 VPAVENTRQVLKRSIRD 89
RESULT 7
PYAS HUMAN
ID PYAS5 HUMAN STANDARD; PRT; 892 AA.
AC P59044;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE PYRIN-containing APAF1-like protein 5.
GN PYPAF5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22162427; PubMed=12019269;
RA Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,
RA Lora J.M., Geddes B.J., Briskin M., DiStefano P.S., Bertin J.;
RT "PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates
activation of NF-kappa B and caspase-1-dependent cytokine
processing";
RL J. Biol. Chem. 277:29874-29880(2002).
RN [2]
RP FUNCTION.
RX MEDLINE=22275822; PubMed=12387869;
RA Grenier J.M., Wang L., Manji G.A., Huang W.-J., Al-Garawi A.,
RA Kelly R., Carlson A., Merriam S., Lora J.M., Briskin M.,
RA DiStefano P.S., Bertin J.;
RT "Functional screening of five PYPAF family members identifies PYPAF5
as a novel regulator of NF-kappaB and caspase-1";
RL FEBS Lett. 530:73-78(2002).
CC -!- FUNCTION: May mediate activation of CASP1 via ASC and promote
activation of NF-kappa-B.
CC -!- SUBUNIT: Binds to ASC with its DAPIN domain.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Highly expressed in granulocytes. Detected at
much lower levels in T-cells.
CC -!- SIMILARITY: Contains 1 DAPIN domain.

Buckingham J.M., Moyzis R.K., Deaven L.L., Doggett N.A.; "Ancient missense mutations in a new member of the Roret gene family are likely to cause familial Mediterranean fever."; Cell 90:797-807(1997).

[2]

RN SEQUENCE OF 305-754 FROM N.A., AND VARIANTS FMF.
RN MEDLINE=97434208; PubMed=9288094;

RA Bernot A., Clepet C., Dasilva C., Devaud C., Petit J.-L.,
RA Caloustian C., Cruaud C., Samson D., Pulcini F., Weissenbach J.,
RA Heilig R., Notaricola C., Domingo C., Rozenbaum M., Bencherit E.,
RA Topaloglu R., Dewalle I., Dross C., Hadjari P., Dupont M.,
RA Demaille J., Toutou I., Smoui N., Nedelec B., Mery J.-P.,
RA Chaabouni H., Delpech M., Grateau G.

RT "A candidate gene for familial Mediterranean fever.";
RL Nat. Genet. 17:25-31(1997).

RN [3]
RNP VARIANTS FMF, AND VARIANT GLN-202.
RX MEDLINE=98334552; PubMed=9668175;
RA Bernot A., da Silva C., Petit J.-L., Cruaud C., Caloustian C.,
RA Castet V., Ahmed-Arab M., Dross C., Dupont M., Cattani D., Smoui N.,
RA Dode C., Pecheux C., Nedelec B., Medaxian J., Rozenbaum M., Rosner I.,
RA Delpech M., Grateau G., Demaille J., Weissenbach J., Toutou I.;
RT "Non-founder mutations in the MFV gene establish this gene as the
cause of familial Mediterranean fever (FMF).";
RL Hum. Mol. Genet. 7:1317-1325(1998).

RN [4]
RNP VARIANTS FMF ILE-680; ILE-681; ILE-694; VAL-694; MET-694 DEL AND
RX ALA-726.
RX MEDLINE=99149053; PubMed=10024914;
RA Booth D.R., Gilmore J.D., Booth S.E., Pepys M.B., Hawkins P.N.;
RT "Pyrimin/mareonstrin mutations in familial Mediterranean fever.";
RL QJM 91:603-606(1998).

RN [5]
RNP VARIANTS FMF.
RX MEDLINE=99192341; PubMed=10090880;
RA Aksentjevich I., Torosyan Y., Samuels J., Centola M., Pras E.,
RA Chae J.J., Oddoux C., Wood G., Azzaro M.P., Palumbo G., Giustolisi R.,
RA Pras M., Ostrer H., Kastner D.L.;
RT "Mutation and haplotype studies of Familial Mediterranean Fever reveal
new ancestral relationships and evidence for a high carrier frequency
with reduced penetrance in the Ashkenazi Jewish population.";
RL Am. J. Hum. Genet. 64:949-962(1999).

RN [6]
RNP VARIANTS FMF GLN-148; SER-369; GLN-408; LEU-479; ILE-680; VAL-694;
RX ALA-726 AND HIS-761.
RX MEDLINE=99294585; PubMed=10364520;
RA Cazeneuve C., Sarkisian T., Pecheux C., Derwichian M., Nedelec B.,
RA Reinert P., Ayvazyan A., Kouymoudjian J.-C., Arapsetan H.,
RA Delpech M., Goossens M., Dode C., Grateau G., Anselem S.;
RT "MEFV-Gene analysis in Armenian patients with Familial Mediterranean
fever: diagnostic value and unfavorable renal prognosis of the M694V
homozygous genotype-genetic and therapeutic implications.";
RL Am. J. Hum. Genet. 65:88-97(1999).

RN [7]
RNP VARIANTS FMF ILE-680; ILE-694; VAL-694 AND ALA-726.
RX MEDLINE=99250763; PubMed=10234504;
RA Shohat M., Magal N., Shohat T., Chen X., Dagan T., Mmouni A.,
RA Danon Y., Lotan R., Ogur G., Sirin A., Schlezinger M., Halpern G.J.,
RA Schwabe A., Kastner D., Rotter J.I., Fischel-Ghodsian N.;
RT "Phenotype-genotype correlation in familial Mediterranean fever:
evidence for an association between Met694Val and amyloidosis.";
RL Eur. J. Hum. Genet. 7:287-292(1999).

RN [8]
RNP VARIANTS FMF GLN-148; ILE-680; ILE-694; ARG-695; ALA-726 AND
RX HIS-761.
RX MEDLINE=20081069; PubMed=10612841;
RA Akar N., Misiroglu M., Walczinka F., Akar E., Cakar N., Tumer N.,
RA Akcaks M., Tastan H., Matcinkeya Y.

RT "MEFV mutations in Turkish patients suffering from Familial
Mediterranean Fever.";
RL Hum. Mutat. 15:118-119(2000).

RN [9]
RNP VARIANT GLN-148.

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF464765; AAL69963.1; -;
 DR InterPro; IPR007091; LRR RNinH.
 DR InterPro; IPR007111; NACHT NTPase.
 DR InterPro; IPR004020; PAAD DAPIN_dom.
 DR Pfam; PF02758; PAAD_DAPIN; 1.
 DR PROSITE; PS0824; DAPIN; 1.
 DR PROSITE; PS50837; NACHT; 1.
 KW ATP-binding; Leucine-rich repeat; Repeat.
 FT DOMAIN 1 93 DAPIN.
 FT DOMAIN 172 491
 FT REPEAT 614 638 LRR 1.
 FT REPEAT 674 697 LRR 2.
 FT REPEAT 760 784 LRR 3.
 FT REPEAT 788 810 LRR 4.
 FT REPEAT 817 840 LRR 5.
 FT REPEAT 845 868 LRR 6.
 FT REPEAT 874 897 LRR 7.
 FT REPEAT 902 928 LRR 8.
 FT REPEAT 933 957 LRR 9.
 FT NP BIND 178 185 ATP (POTENTIAL).
 SQ SEQUENCE 980 AA; 111806 MW; 822AF2FD4338003D CRC64;
 Query Match 10.1%; Score 99.5; DB 1; Length 980;
 Best Local Similarity 26.6%; Pred. No. 0.71;
 Matches 45; Conservative 21; Mismatches 80; Indels 23; Gaps 5;
 QY 12 LNLTAELKPKLLSVPLREGYGRIPRGALLSMDALDLDTKLVSFYLETYGAEALTAN 71
 Db 14 LQQLNEDELLKPSKLLWAPLEEDVLOKTPWSEVADGKKAEILVNTSSNWRNATVN 73
 QY 72 VLRLDGLQEMA-----GQLQAATHQSGGAAPAGIQAPPSAAKFLGH---FIDGH 118
 Db 74 ILEENMLTELCKMAKEMMEDQVQIDNPDLGDA-----EEDSELAKFGEKGEGRNSWE 128
 QY 119 RAALITRVNTEWLLDALYKGLVTBQYQVRAEP-TNPSKVRKLFST 166
 Db 129 KQSLVYKNTFWQGDIDNFHDVTLRNQ----RFIFPLNPTPRKLTPT 173
 RESULT 10
 CAR8 HUMAN
 ID CAR8 HUMAN STANDARD; PRT; 431 AA.
 AC Q9Y2G2; Q96P82;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Caspase recruitment domain protein 8 (Apoptotic protein NDPPI) (DACAR)
 DE (CARD-inhibitor of NF-kappaB activating ligand) (CARDINAL) (TUCAN).
 GN CARD8 OR NDPPI OR KIAA0955.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM LONG).
 RC TISSUE=Brain;
 RX MEDLINE=99246063; PubMed=10231032;
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosewa M.,
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XIII.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 6:63-70(1999).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM LONG).
 RX MEDLINE=21950691; PubMed=11821383;

RA Razmara M., Srinivasula S.M., Wang L., Poyet J.-L., Geddes B.J.,
 RA DiStefano P.S., Berlin J., Alnemri E.S.;
 RT "CARD-8 protein, a new CARD family member that regulates caspase-1
 RT activation and apoptosis.";
 RL J. Biol. Chem. 277:13952-13958(2002).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM LONG).
 RA Zhang H.;
 RT "A novel apoptotic protein, NDPPI, containing CARD and BH3 domains.";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM LONG).
 RA Guet C., Vito P.;
 RT "DACAR, a novel CARD-containing protein.";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM LONG).
 RX MEDLINE=21570185; PubMed=11551959;
 RA Bouchier-Hayes L., Conroy H., Egan H., Adrain C., Creagh E.M.,
 RA MacFarlane M., Martin S.J.;
 RT "CARDINAL, a novel caspase recruitment domain protein, is an inhibitor
 RT of multiple NF-kappa B activation pathways.";
 RL J. Biol. Chem. 276:44069-44077(2001).
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM SHORT).
 RC TISSUE=Kidney;
 RA Guo J.H., Yu L.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP CHARACTERIZATION.
 RX MEDLINE=21402909; PubMed=11408476;
 RA Pathan N., Marusawa H., Krajewska M., Matsuzawa S.-I., Kim H.,
 RA Okada K., Torii S., Kitada S., Krajewski S., Welsh K., Pio F.,
 RA Godzik A., Reed J.C.;
 RT "TUCAN, an antiapoptotic caspase-associated recruitment domain family
 RT protein overexpressed in cancer.";
 RL J. Biol. Chem. 276:32220-32229(2001).
 RN [8]
 RP CHARACTERIZATION, AND MUTAGENESIS OF LYS-366.
 RX MEDLINE=22062958; PubMed=12067710;
 RA Stilo R., Leonardi A., Formisano L., Di Jeso B., Vito P., Liguoro D.;
 RT "TUCAN/CARDINAL and DRAL participate in a common pathway for
 RT modulation of NF-kappaB activation.";
 RL FEBS Lett. 521:165-169(2002).
 CC -!- FUNCTION: Inhibits NF-kappaB activation. May participate in a
 CC regulatory mechanism that coordinates cellular responses
 CC controlled by NF-kappaB transcription factor. Involved in the
 CC negative regulation of caspase-1.
 CC -!- SUBUNIT: May form homodimers. Interacts with NEMO and DRAL. Binds
 CC to caspase-1, pseudo-ICE and ICEBERG.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Name=Long;
 CC IsoId=Q9Y2G2-1; Sequence=Displayed;
 CC Name=Short;
 CC IsoId=Q9Y2G2-2; Sequence=VSP_000782; VSP_000783;
 CC -!- TISSUE SPECIFICITY: High expression in lung, ovary, testis and
 CC placenta. Lower expression in heart, kidney and liver. Also
 CC expressed in spleen, lymph node and bone marrow.
 CC -!- SIMILARITY: Contains 1 CARD domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB023172; BAA76799.1; -;
 DR EMBL; AF322184; AAG50014.1; -;
 DR EMBL; AF331519; AAK01126.1; -;

```

DR EMBL; AY026322; AAK09882.1; -
DR EMBL; AF405558; AAL02427.1; -
DR EMBL; AF511652; AAM46959.1; -
DR InterPro; IPR001315; CARD.
DR Pfam; PF006119; CARD; 1.
DR SMART; SM00114; CARD; 1.
DR PROSITE; PS0209; CARD; 1.
KW Apoptosis; Nuclear protein; Alternative splicing.
FT DOMAIN 340 430
FT VARSPLIC 282 286
FT VARSPLIC 287 431
FT MUTAGEN 366 431
FT CONFLICT 366 60
FT CONFLICT 326 326
FT CONFLICT 422 422
FT CONFLICT 422 422
SQ SEQUENCE 431 AA; 48932 MW; CB54D130807732E6 CRC64;

Query Match
Best Local Similarity 24.0%; Score 99; DB 1; Length 431;
Matches 42; Conservative 32; Mismatches 73; Indels 28; Gaps 5;

QY 21 KFKLKLKLVLPREGYGRIPRGALLSMDALDITDKLVSPYLETYGAELTANVLRDMGLQE 80
DB 281 KELKLSYRSPGEIQHFSKVFAGQMKPEIQLEITEK-----RHGTLVMDTEVKPVDLQL 333
QY 81 MAGQLQAATHQSGAAGAPAGIQAPQQAAGPLHFDIHOHRAALIAVTVNVEMLLDALY-CK 139
DB 334 VA-----ASAP-----PPFS-----GAAFVKNHRQLQARMDGLKGLVDLDLQDNE 373
QY 140 VLTDEQYQVRAEPTNPSPKMKLFSTPAWNWTKDLQLLQALRESQSYLVEDLER 194
DB 374 VLTENEKELVEQEKTRQSKNEALLSVNVEKGLDALDVLFRSISERDPYLVSVLRQ 428

RESULT 11
SIDE_SALT
ID _SIDE_SALT STANDARD; PRT; 143 AA.
AC Q9ZAN3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Silver-binding protein silp precursor.
GN SILE
OS Salmonella typhimurium.
OG Plasmid pMG101.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]_
RP SEQUENCE FROM N.A., SEQUENCE OF 21-35, AND CHARACTERIZATION.
RX MEDLINE=99128056; PubMed=9930866;
RA Gupta A., Matsui K., Lo J.-F., Silver S.;
RT "Molecular basis for resistance to silver cations in Salmonella.";
RL Nat. Med. 5:183-188(1999).
CC -!- FUNCTION: COMPONENT OF THE SIL CATION-EFFLUX SYSTEM THAT CONFERS
CC -!- RESISTANCE TO SILVER.
CC -!- SUBCELLULAR LOCATION: Periplasmic.
CC -!- INDUCTION: By silver.
CC -!- SIMILARITY: TO E-COLI PCOE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF067954; AAD11743.1; -
KW Metal-binding; periplasmic; Signal; Plasmid.
FT SIGNAL 1 20

```

```

FT CHAIN 21 143 SILVER-BINDING PROTEIN SILE.
SQ SEQUENCE 143 AA; 15201 MW; 703B9BC0F8BDOCB6 CRC64;

Query Match
Best Local Similarity 9.2%; Score 91.5; DB 1; Length 143;
Matches 31; Conservative 11; Mismatches 47; Indels 9; Gaps 4;

QY 64 YGAELTANVLRDMGLQEMAGQLQAATHQ-GSGAAGAPAGIQAPQQAAGPLHFDIHOHRAAL 122
DB 12 FGLSSAWATETVNIHERVNNQAQAHAHQMSAAAPVGIQ-----GTAPRMAGMDQHEQAI 66
QY 123 IARVTNVEMLLDALYGVKVLTDQYQ--AVRAEPTNPSPK 158
DB 67 IAHETMTNGSADA-HQKMWESHQRMWGSQTVSPGSPK 103

RESULT 12
RUVB_THETN
ID RUVB_THETN STANDARD; PRT; 338 AA.
AC Q8RAN2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Holliday junction DNA helicase ruvB.
GN RUVB OR TTE1180.
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
CC -!- FUNCTION: The ruvA-ruvB complex in the presence of ATP renatures
CC cruciform structure in supercoiled DNA with palindromic sequence,
CC indicating that it may promote strand exchange reactions in
CC homologous recombination. RuvAB is an helicase that mediates the
CC Holliday junction migration by localized denaturation and
CC reannealing (By similarity).
CC -!- SUBUNIT: Forms a complex with ruvA (By similarity).
CC -!- SIMILARITY: BELONGS TO THE RUVB FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE013080; AAM24411.1; -
DR HAMAP; MF_00016; -; 1.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003959; AAA_ATPase_cent.
DR InterPro; IPR004605; RuvB.
DR Pfam; PF00004; AAA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR00635; ruvB; 1.
KW DNA repair; SOS response; ATP-binding; DNA recombination; Helicase;
KW Complete proteome.
FT NP_BIND 59 66 ATP (POTENTIAL).
FT SEQUENCE 338 AA; 38072 MW; 2B9562D8CBD988D1 CRC64;

Query Match
Best Local Similarity 9.0%; Score 89; DB 1; Length 338;
Matches 39; Conservative 32; Mismatches 73; Indels 18; Gaps 5;

QY 3 RARDAILDALENLTAEELK---KFKLKLKLSVPLREG-----YGRIPRGALLSMDAL----- 50

```



```
Db 169 RSRFGVNRDYSVEELKEIKRSANILNIGIDEADAFAEIAARRSGTGPRIANRLKRV 228
Qy 51 DLTDKLVSLTYGAELTANVL--RDMGLQWAGLOQAATHQSGAAPAGIOAPPQSA 108
Db 229 DPAEVKNGYIDYNTANTALNMLGVDEMGLEIDRKILIAIEKEFGGGVGVGIDATA 288
Qy 109 KPLHFDIOHRAALARTVNVEWMLDALYKVLTDQVQAVR 150
Db 289 EDGTIEDWVEPYLM---QIGFLNRTPRGRVVTIKLAYQYLK 326

RESULT 13
M22_STRPY STANDARD; PRT; 372 AA.
AC P50569;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE M protein, serotype 2.2 precursor.
GN EML2.2.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T2/44/RB4/119;
RX MEDLINE=92104662; PubMed=1370269;
RA Bessen D.E., Fischetti V.A.;
RT "Nucleotide sequences of two adjacent M or M-like protein genes of
RT group A streptococci: different RNA transcript levels and
RT identification of a unique immunoglobulin A-binding protein.";
RL Infect. Immun. 60:124-135(1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Fischetti V.A., Bessen D.E.;
RT "Immunoglobulin A binding protein.";
RL Patent number US556944, 17-SEP-1996.
CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE DIFFERENT ANTIGENIC SEROTYPES
CC OF PROTEIN M. PROTEIN M IS CLOSELY ASSOCIATED WITH VIRULENCE OF
CC THE BACTERIUM AND CAN RENDER THE ORGANISM RESISTANT TO
CC PHAGOCYTOSIS.
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (Potential).
CC -!- SIMILARITY: TO OTHER M PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X61276; CAA43582.1; -
CC EMBL; I26204; -; NOT_ANNOTATED_CDS.
CC PIR; S23326; S23326.
CC InterPro; IPR005877; Gpos_Y5IRK.
CC InterPro; IPR001899; Gram_pos_anchor.
CC InterPro; IPR006192; LPXTG.
CC InterPro; IPR003345; M_repeat.
CC Pfam; PF00746; Gram_pos_anchor; 1.
CC Pfam; PF02370; M; 8.
CC Pfam; PF04650; Y5IRK_signal; 1.
CC PRINTS; PR00015; GPOSANCHOR.
CC TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
CC TIGRFAMs; TIGR01168; Y5IRK_signal; 1.
CC PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
KW Virulence; Phagocytosis; Cell wall; Peptidoglycan-anchor; Repeat;
KW Antigen; Coiled coil; Signal.
FT SIGNAL 1 41 POTENTIAL.
FT CHAIN 42 342 M PROTEIN, SEROTYPE 2.2.
```

```
FT PROPEP 343 372 REMOVED BY SORTASE (POTENTIAL).
FT DOMAIN 131 244 3 X REPEATS, TYPE C.
FT REPEAT 131 153 C-1.
FT REPEAT 173 195 C-2.
FT REPEAT 222 244 C-3.
FT DOMAIN 305 338 GLY/PRO-RICH (CELL WALL-SPANNING)
FT SITE 339 343 (BY SIMILARITY).
FT MOD RES 342 LPXTG SORTING SIGNAL (POTENTIAL).
SQ SEQUENCE 372 AA; 41149 MW; E8FD5D0920C95C74 CRC64;
      AMIDE-LINKED TO CELL WALL (POTENTIAL).

Query Match      8.7%; Score 86.5; DB 1; Length 372;
Best Local Similarity 26.8%; Pred. No. 3.1;
Matches 53; Conservative 30; Mismatches 78; Indels 37; Gaps 9;

Qy 12 LENTAEELKKFKLLSLVPLREGYGRIPRGALLSMDAL-----DLTDKLVSYLETYGA 66
Db 78 LEKINAEKEKKKLEAINKELNENYKLDG-----IDALEKEKEDLTKTLAKTTKENEL 133
Qy 67 ELTANVL-RDMGLQWAGLOQAATHQSGAAPAGIOAPPQ--SAAKPGL-HFDIOHRAAL 122
Db 134 EASRKGLSRDLEASRTAKKELEAKHOKLEAENKKLTEGNOVSEASRKGLSNDLEASRAA- 192
Qy 123 IARVTNVEWMLDALYKVLTDQ-----YQAVRAPETNPSPKMKLFSFTPAWNWTKDLL 177
Db 193 -----KKELEAKYQKLETDHQALEAKHOKLEADYQVSETSRKGLS-----RD-- 234
Qy 178 LQALRESQSYLVEDLERS 195
Db 235 LEASREANKKVTSELTQA 252

RESULT 14
SYEB_CORGL STANDARD; PRT; 835 AA.
ID SYEB_CORGL STANDARD; PRT; 835 AA.
AC Q8NQNG;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20)
DE (Phenylalanyl-tRNA ligase beta chain) (PHERS).
GN PHET OR CGI1390.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA(Phe) = AMP +
CC diphosphate + L-phenylalanyl-tRNA(Phe).
CC -!- COFACTOR: Binds 2 magnesium ions per tetramer (By similarity).
CC -!- SUBUNIT: Tetramer of two alpha and two beta chains (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the phenylalanyl-tRNA synthetase beta chain
CC family. Subfamily 1.
CC -!- SIMILARITY: Contains 1 tRNA-binding domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP005278; BAB98783.1; ALT_INIT.
CC HAMAP; MF_00283; -; 1.
CC InterPro; IPR005146; B3_4.
```

```
DR InterPro; IPR005147; B5.
DR InterPro; IPR005121; Pdx-AntiCB.
DR InterPro; IPR004532; PheT bact.
DR InterPro; IPR002547; tRNA_bind.
DR Pfam; PF03483; B3_4; 1.
DR Pfam; PF03484; B5; 1.
DR Pfam; PF03147; FDX-ACB; 1.
DR Pfam; PF01588; tRNA_bind; 1.
DR TIGRfam; TIGR00472; pHeT_bact; 2.
DR PROSITE; PS50886; TRBD; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Metal-binding; Magnesium; RNA-binding; tRNA-binding;
KW Complete proteome.
FT DOMAIN 44 160 TRNA-BINDING.
FT METAL 472 472 MAGNESIUM (BY SIMILARITY)
FT METAL 478 478 MAGNESIUM (VIA CARBONYL OXYGEN) (BY
SIMILARITY).
FT METAL 481 481 MAGNESIUM (BY SIMILARITY).
FT METAL 482 482 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 835 AA; 89435 MW; 285EC9A0673DA49F CRC64;

Query Match 8.7%; Score 86.5; DB 1; Length 835;
Best Local Similarity 24.5%; Pred. No. 7.9;
Matches 40; Conservative 29; Mismatches 73; Indels 21; Gaps 7;

QY 29 SVPLREGYGRIPRGALLSMDALDLDTKLVSYL---ETYGAEALTANVLDRMDGLQEWAGOL 85
Db 114 ASARTYGRNAGMICSASELGLADKQNSGITLDPSTG-FPGEDARQALGUEDTVFVDV 172
QY 86 QAATHQSGAAPAGIQAPQSAKPGLFHFDHRAALIARVNVEVLLDALYGVKLT--- 142
Db 173 NVTDPDRGVALSARGLTRELASAF--SLTFTD---PAIEPAVAGIEVKVPAVEGSLINVEL 227
QY 143 DRQYQAVRAEPTNPSSKMKLFSTPAWN---WTCKDLLLQALR 182
Db 228 REETKAIR-----FGLRKVSGIDPAESPFPMQRELMSGQR 264

RESULT 15
CIS1_MOUSE STANDARD; PRT; 1033 AA.
AC Q8R4B8;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cold autoinflammatory syndrome 1 protein homolog (PYRIN-containing
DE APAF1-like protein 1) (Mast cell maturation inducible protein 1).
GN CIA1 OR PYPAF1 OR MMIG1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/cJ;
RA Kikuchi-Yanoshita R., Koga K., Taketomi Y., Sugiki T., Saito T.,
RA Ishii S., Hisada M., Suzuki-Nishimura T., Uchida M.K., Moon T.-C.,
RA Chang H.-W., Sawada M., Inagaki N., Nagai H., Murakami M., Kudo I.;
RT Identification of inducible genes during in vitro maturation of mouse
RT bone marrow-derived mast cells to connective tissue-type mast cells."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May function as a potential inducer of apoptosis.
CC Interacts selectively with apoptosis-associated specklike protein
CC containing a CARD domain (ASC). This complex may function as an
CC upstream activator of NF-kappaB signaling (By similarity).
CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

```
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF486632; AAL90874.1; -.
DR MGD; MGI:2653833; Cla1.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR003590; LRR_RNinh_sub.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF02758; PAAD_DAPIN; 1.
DR SMART; SM00368; LRR_R1; 1.
DR PROSITE; PS50824; DAPIN; 1.
DR PROSITE; PS50837; NACHT; 1.
KW Apoptosis; Repeat; Leucine-rich repeat.
FT DOMAIN 1 91 DAPIN.
FT DOMAIN 216 532 NACHT.
FT REPEAT 737 760 LRR 1.
FT REPEAT 794 817 LRR 2.
FT REPEAT 851 874 LRR 3.
FT REPEAT 880 903 LRR 4.
FT REPEAT 908 931 LRR 5.
FT REPEAT 937 964 LRR 6.
FT REPEAT 965 988 LRR 7.
SQ SEQUENCE 1033 AA; 118274 MW; 5924690966B12117 CRC64;

Query Match 8.7%; Score 86; DB 1; Length 1033;
Best Local Similarity 35.6%; Pred. No. 11;
Matches 21; Conservative 9; Mismatches 29; Indels 0; Gaps 0;

QY 1 MGRARDATLDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLDTKLVSF 59
Db 1 MTSVRCKLAQVLEDDVDLKKFKMHLEDYPPKGCIPVPRGQMEKADHLDLATLIMDF 59

Search completed: January 29, 2004, 13:46:32
Job time : 5.00185 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 13:06:01 ; Search time 10.8067 Seconds
(without alignments)
4656.416 Million cell updates/sec

Title: US-09-996-617-8

Perfect score: 990

Sequence: 1 MGRARDAILDALENLTAEEL.....LLQLALRSQSYLVEDLERS 195

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPREMBL_23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_podent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	736	74.3	195	6 Q8HXK9	Q8hxk9 bos taurus
2	673	68.0	193	11 Q8CHK8	Q8chk8 rattus norv
3	275	27.8	89	4 Q8NFP8	Q8nfp8 homo sapien
4	267	27.0	89	4 Q8WXC3	Q8wxc3 homo sapien
5	142	14.3	750	11 Q9UJ25	Q9ij25 rattus norv
6	142	14.3	767	11 Q9UJ26	Q9ij26 mus musculus
7	104	10.5	404	13 Q9DDJ2	Q9ddj2 brachydanio
8	91.5	9.2	127	2 Q8VPY3	Q8vpy3 serratia ma
9	91.5	9.2	127	2 Q8VPW8	Q8vpw8 salmonella
10	91.5	9.2	127	2 Q8VPX8	Q8vpx8 salmonella
11	91.5	9.2	127	2 Q8VPY8	Q8vpy8 serratia ma
12	91.5	9.2	383	13 Q9JL7	Q9j9l7 brachydanio
13	86.5	8.7	439	16 P74033	P74033 synecocyst
14	86.5	8.7	828	16 Q8NQ6	Q8nq6 corynebacte
15	86	8.7	580	17 Q8TSR7	Q8tsr7 methanosarc
16	86	8.7	1033	11 Q8R4B8	Q8r4b8 mus musculus

ALIGNMENTS

RESULT 1

ID	Q8HXK9	PRELIMINARY;	PRT;	195 AA.
AC	Q8HXK9;			
DT	01-MAR-2003 (Tremblrel. 23, Created)			
DT	01-MAR-2003 (Tremblrel. 23, Last sequence update)			
DT	01-MAR-2003 (Tremblrel. 23, Last annotation update)			
DE	Apoptosis-associated speck-like protein containing a CARD.			
GN	BASC.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI_TaxID=9913;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RP	Masumoto J., Zhou W., Chen P.F., Su F., Kuwada J.Y., Hidaka E.,			
RA	Katsuyama T., Sagara J., Taniguchi S., Ngo-Hazelett P.,			
RA	Postlethwait J.H., Nunez G., Inohara N.,			
RT	"Casp-1: A zebrafish caspase activated by ASC oligomerization required for pharyngeal Arch development."			
RL	J. Biol. Chem. 0:0-0(2002).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20036508; PubMed=10567338;			
RA	Masumoto J., Taniguchi S., Ayukawa K., Sarvatham H., Kishino T.,			
RA	Niikawa N., Hidaka E., Katsuyama T., Higuchi T., Sagara J.;			
RT	"ASC, a novel 22-kDa protein, aggregates during apoptosis of human promyelocytic leukemia HL-60 cells;"			
RL	J. Biol. Chem. 274:33835-33838(1999).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20580347; PubMed=11139337;			
RA	Masumoto J., Taniguchi S., Nakayama K., Ayukawa K., Sagara J.;			
RT	"Murine ortholog of ASC, a CARD-containing protein, self-associates, and exhibits restricted distribution in developing mouse embryos."			
RL	Exp. Cell Res. 262:128-133(2001).			
DR	EMBL; AB050006; BAC43753.1; "			
SQ	SEQUENCE 195 AA; 21917 MW; 7C9D4BD8DBA9A9E8 CRC64;			

17	85	8.6	1488	16	Q8Z7Z5	Q8z7z5 salmonella
18	85	8.6	1488	16	Q935S7	Q935s7 salmonella
19	84.5	8.5	894	10	Q8GSF4	Q8gsf4 oryza sativ
20	84.5	8.5	1241	3	Q9HEM2	Q9hem2 neurospora
21	84	8.5	292	16	Q25976	Q25976 helicobacte
22	84	8.5	530	5	Q8T9T6	Q8t9t6 aedes aegypt
23	83.5	8.4	542	17	Q9HRG1	Q9hrg1 halobacteri
24	83.5	8.4	2556	12	Q9QH56	Q9qh56 gallid herp
25	83	8.4	469	16	Q8CVN9	Q8cvn9 escherichia
26	83	8.4	545	16	Q9CBL2	Q9cbl2 mycobacteri
27	83	8.4	877	2	Q45095	Q45095 bacillus ci
28	82.5	8.3	409	5	Q96984	Q96984 stylonychia
29	82.5	8.3	522	2	Q85210	Q85210 vibrio chol
30	82.5	8.3	522	16	Q9K394	Q9k394 vibrio chol
31	82.5	8.3	567	3	Q8NIK9	Q8nik9 cryptococcu
32	82	8.3	292	16	Q9ZJ11	Q9zj11 helicobacte
33	81	8.2	456	4	Q9BUC6	Q9buc6 homo sapien
34	81	8.2	503	2	Q9XEW8	Q9xew8 rhodobacter
35	81	8.2	631	4	Q8NDM4	Q8ndm4 homo sapien
36	81	8.2	631	4	Q969Z0	Q969z0 homo sapien
37	81	8.2	792	16	Q8ZPS3	Q8zps3 salmonella
38	81	8.2	792	16	Q8ZGJ0	Q8zgj0 salmonella
39	81	8.2	896	2	Q9AN79	Q9an79 bradyrhizob
40	81	8.2	9477	2	Q9L4X3	Q9l4x3 streptomyc
41	80.5	8.1	266	10	Q8LM81	Q8lm81 oryza sativ
42	80.5	8.1	718	16	Q9JYB6	Q9jyb6 neisseria m
43	80.5	8.1	725	16	Q9JT97	Q9jty97 neisseria m
44	80.5	8.1	727	16	Q8DE24	Q8de24 vibrio vuln
45	80.5	8.1	972	4	Q8N612	Q8n612 homo sapien

```
Query Match      74.3%; Score 736; DB 6; Length 195;
Best Local Similarity 75.4%; Pred. No. 1.7e-59;
Matches 147; Conservative 20; Mismatches 28; Indels 0; Gaps 0;

QY 1 MGRDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60
DB 1 MGTTRDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSY 60
QY 61 LETYGAELTANVLRDMGLQEMAGLOQAATHQSGAPAGIQAPQSAAPKPGHLHFIDQHRA 120
DB 61 LEAYGAELTANVLRDMGLQEMAGLOQAATHQSGAPAGIQAPQSAAPKPGHLHFIDQHRA 120
QY 121 ALIARVTNVVWLLDALYGVKLVLTDEQYQAVRAEPTNPSKMKLFSPFPAWNTCKOLLLOA 180
DB 121 ALIARVTNVVWLLDALYGVKLVLTDEQYQAVRAEPTNPSKMKLFSPFPAWNTCKOLLLOA 180
QY 181 LRESQSYLVEDLERS 195
DB 181 LRDTQPYLVDDLEQS 195

RESULT 2
Q8CHK8 PRELIMINARY; PRT; 193 AA.
AC Q8CHK8;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Apoptosis-associated speck-like protein.
GN RASC.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Masumoto J., Zhou W., Chen F.F., Su F., Kuwada J.Y., Hidaka E.,
RA Katsuyama T., Sagara J., Taniguchi S., Ngo-Hazlett P.,
RA Postlethwait J.H., Nunez G., Inohara N.,
RT "Casp9: A zebrafish caspase activated by ASC oligomerization required
RT for pharyngeal Arch development.";
RL J. Biol. Chem. 0:0-0(2002).
DR EMBL; AB053165; BAC43754.1; -.
SQ SEQUENCE 193 AA; 21654 MW; F3B27B560D86A17B CRC64;

Query Match      68.0%; Score 673; DB 11; Length 193;
Best Local Similarity 69.7%; Pred. No. 9.7e-54;
Matches 136; Conservative 20; Mismatches 37; Indels 2; Gaps 2;

QY 1 MGRDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60
DB 1 MGTTRDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSY 60
QY 61 LETYGAELTANVLRDMGLQEMAGLOQAATHQSGAPAGIQAPQSAAPKPGHLHFIDQHRA 120
DB 61 LRGYGLELTANVLRDMGLQEMAGLOQAATHQSGAPAGIQAPQSAAPKPGHLHFIDQHRA 118
QY 121 ALIARVTNVVWLLDALYGVKLVLTDEQYQAVRAEPTNPSKMKLFSPFPAWNTCKOLLLOA 180
DB 119 ALIARVTNVVWLLDALYGVKLVLTDEQYQAVRAEPTNPSKMKLFSPFPAWNTCKOLLLOA 178
QY 181 LRESQSYLVEDLERS 195
DB 179 LRQTOPYLVTDLEQS 193

RESULT 3
Q8NFP8 PRELIMINARY; PRT; 89 AA.
AC Q8NFP8;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)

Query Match      74.3%; Score 736; DB 6; Length 195;
Best Local Similarity 75.4%; Pred. No. 1.7e-59;
Matches 147; Conservative 20; Mismatches 28; Indels 0; Gaps 0;

QY 1 MGRDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60
DB 1 MGTTRDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSY 60
QY 61 LETYGAELTANVLRDMGLQEMAGLOQAATHQSGAPAGIQAPQSAAPKPGHLHFIDQHRA 120
DB 61 LEAYGAELTANVLRDMGLQEMAGLOQAATHQSGAPAGIQAPQSAAPKPGHLHFIDQHRA 120
QY 121 ALIARVTNVVWLLDALYGVKLVLTDEQYQAVRAEPTNPSKMKLFSPFPAWNTCKOLLLOA 180
DB 121 ALIARVTNVVWLLDALYGVKLVLTDEQYQAVRAEPTNPSKMKLFSPFPAWNTCKOLLLOA 180
QY 181 LRESQSYLVEDLERS 195
DB 181 LRDTQPYLVDDLEQS 195

RESULT 4
Q8WXC3 PRELIMINARY; PRT; 89 AA.
AC Q8WXC3;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Pyrin-only protein 1.
GN POPI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bertin J.;
RT "POPI: a pyrin-only protein that regulates inflammatory signaling.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF454669; AAL58439.1; -.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF02758; PAAD_DAPIN; 1.
DR PROSITE; PS50824; DAPIN; 1.
SQ SEQUENCE 89 AA; 10107 MW; 4CDF6D672DDDD98E CRC64;

Query Match      27.0%; Score 267; DB 4; Length 89;
Best Local Similarity 64.8%; Pred. No. 5.6e-17;
Matches 57; Conservative 9; Mismatches 22; Indels 0; Gaps 0;

QY 1 MGRDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60
DB 1 MGTTRDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSY 60
QY 61 LETYGAELTANVLRDMGLQEMAGLOQA 88
DB 61 YEDYAAELVAVLRDMRMLLEEAARLQRA 88

RESULT 5
Q8JUJ25 PRELIMINARY; PRT; 750 AA.
AC Q8JUJ25;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
```

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 GN Pyrin (Marenostarin).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RC STRAIN=Sprague-Dawley;
 RX MEDLINE=20279845; PubMed=10818206;
 RA Chae J.J., Centola M., Aksentijevich I., Dutra A., Tran M., Wood G.,
 RA Nagaraju K., Kingma D.W., Liu P.P., Kastner D.L.;
 RT "Isolation, genomic organization, and expression analysis of the mouse
 RT and rat homologs of MEV, the gene for familial Mediterranean fever.";
 RL Mamm. Genome 11:428-435(2000).
 CC -!- FUNCTION: PROBABLY CONTROLS THE INFLAMMATORY RESPONSE IN
 CC MYELOMONOCYTIC CELLS AT THE LEVEL OF THE CYTOSKELETON ORGANIZATION
 CC (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH MICROTUBULES
 CC AND WITH THE FILAMENTOUS ACTIN OF PERINUCLEAR FILAMENTS AND
 CC PERIPHERAL LAMELLAR RUFFLES (BY SIMILARITY).
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN SPLEEN AND, TO A LESSER DEGREE IN
 CC THE LUNG. NOT EXPRESSED IN THYMUS, TESTIS, OVARY, HEART, BRAIN,
 CC LIVER, KIDNEY AND MUSCLE.
 CC -!- SIMILARITY: CONTAINS 1 B BOX-TYPE ZINC FINGER.
 CC -!- SIMILARITY: CONTAINS 1 DAPIN DOMAIN.
 DR EMBL; AF143410; AAF03767.1; -
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR Pfam; IPR000315; Znf_Bbox.
 DR Pfam; PF02758; PAAD_DAPIN; 1.
 DR Pfam; PF00643; zf-B_box; 1.
 DR SMART; SM00336; BBOX; 1.
 DR PROSITE; PS50824; DAPIN; 1.
 DR PROSITE; PS50119; ZF_BBOX; 1.
 DR Inflammatory response; Actin-binding; Metal-binding; Cytoskeleton;
 KW Microtubules; Zinc-finger; Zinc.
 FT DOMAIN 1 92 DAPIN.
 FT ZN FING 440 482 B BOX-TYPE.
 FT SEQUENCE 750 AA; 83994 MW; C76D0F3E02711312 CRC64;
 Query Match 14.3%; Score 142; DB 11; Length 750;
 Best Local Similarity 33.7%; Pred. No. 0.00029;
 Matches 35; Conservative 19; Mismatches 50; Indels 0; Gaps 0;
 QY 6 DAILEDLENLTAEELKFKLLSVLPREGYGRIPRGALLSMDALDLDTKLVSYLETYG 65
 Db 7 DHLNTLEELLPYELEKFKFKLHTTSLKGGHSRIPSLVKMARPIKLTLLTLTYGEEYA 66
 QY 66 AELTANVLDMGLQEMAGLOAQATHQGSGAAPAGIQAPPOSAAK 109
 Db 67 VRLTQLILRATNQRQALAEHLKATGHEHLTEENGVSQSSAE 110
 Query Match 14.3%; Score 142; DB 11; Length 750;
 Best Local Similarity 33.7%; Pred. No. 0.00029;
 Matches 35; Conservative 19; Mismatches 50; Indels 0; Gaps 0;
 QY 6 DAILEDLENLTAEELKFKLLSVLPREGYGRIPRGALLSMDALDLDTKLVSYLETYG 65
 Db 7 DHLNTLEELLPYELEKFKFKLHTTSLKGGHSRIPSLVKMARPIKLTLLTLTYGEEYA 66
 QY 66 AELTANVLDMGLQEMAGLOAQATHQGSGAAPAGIQAPPOSAAK 109
 Db 67 VRLTQLILRATNQRQALAEHLKATGHEHLTEENGVSQSSAE 110
 RESULT 6
 Q9JJ26 PRELIMINARY; PRT; 767 AA.
 AC Q9JJ26
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Pyrin (Marenostarin).
 GN MEV.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RC STRAIN=129/Sv;
 RX MEDLINE=20279845; PubMed=10818206;
 RA Chae J.J., Centola M., Aksentijevich I., Dutra A., Tran M., Wood G.,

RA Nagaraju K., Kingma D.W., Liu P.P., Kastner D.L.;
 RT "Isolation, genomic organization, and expression analysis of the mouse
 RT and rat homologs of MEV, the gene for familial Mediterranean fever.";
 RL Mamm. Genome 11:428-435(2000).
 CC -!- FUNCTION: PROBABLY CONTROLS THE INFLAMMATORY RESPONSE IN
 CC MYELOMONOCYTIC CELLS AT THE LEVEL OF THE CYTOSKELETON ORGANIZATION
 CC (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH MICROTUBULES
 CC AND WITH THE FILAMENTOUS ACTIN OF PERINUCLEAR FILAMENTS AND
 CC PERIPHERAL LAMELLAR RUFFLES (BY SIMILARITY).
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN SPLEEN PERIPHERAL BLOOD
 CC GRANULOCYTES. NOT EXPRESSED IN LYMPHOCYTES, THYMUS, TESTIS, OVARY,
 CC HEART, BRAIN, LUNG, LIVER, KIDNEY AND MUSCLE.
 CC -!- SIMILARITY: CONTAINS 1 B BOX-TYPE ZINC FINGER.
 CC -!- SIMILARITY: CONTAINS 1 DAPIN DOMAIN.
 DR EMBL; AF143409; AAF03766.1; -
 DR MGI; MGI:1859396; Mefv.
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR InterPro; IPR000315; Znf_Bbox.
 DR Pfam; PF02758; PAAD_DAPIN; 1.
 DR Pfam; PF00643; zf-B_box; 1.
 DR SMART; SM00336; BBOX; 1.
 DR PROSITE; PS50824; DAPIN; 1.
 DR PROSITE; PS50119; ZF_BBOX; 1.
 DR Inflammatory response; Actin-binding; Metal-binding; Cytoskeleton;
 KW Microtubules; Zinc-finger; Zinc.
 FT DOMAIN 1 92 DAPIN.
 FT ZN FING 439 481 B BOX-TYPE.
 FT SEQUENCE 767 AA; 86264 MW; FF102CB3FD7C1EB CRC64;
 Query Match 14.3%; Score 142; DB 11; Length 767;
 Best Local Similarity 28.3%; Pred. No. 0.0003;
 Matches 32; Conservative 26; Mismatches 47; Indels 8; Gaps 1;
 QY 6 DAILEDLENLTAEELKFKLLSVLPREGYGRIPRGALLSMDALDLDTKLVSYLETYG 65
 Db 7 DHLNTLEELLPYDEKFKFKLQNTSLKGGHSKIPRGHMQRMPVKLASLLTYGEEYA 66
 QY 66 AELTANVLDMGLQEMAGLOAQATHQGSGAAPAGIQAPPOSAAK 110
 Db 67 VRLTQLILRATNQRQALAEHLKATGHEHLTEENGVSQSSVENKAKSVKVP 119
 RESULT 7
 Q9DDJ2 PRELIMINARY; PRT; 404 AA.
 AC Q9DDJ2
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Casp2.
 GN CASP2 OR CASPY2.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Osteiophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20373792; PubMed=10917738;
 RA Inohara N., Nunez G.,
 RT "Genes with homology to mammalian apoptosis regulators identified in
 RT zebrafish.";
 RL Cell Death Differ. 7:509-510(2000).
 DR EMBL; AF327410; AAG45230.1; -
 DR HSP; P29466; IICE.
 DR ZFIN; ZDB-GENE-020812-1; caspb.
 DR InterPro; IPR002398; ICE.
 DR InterPro; IPR002138; ICE_p10.
 DR InterPro; IPR001309; ICE_p20.
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR Pfam; PF00655; ICE_p10; 1.
 DR Pfam; PF00656; ICE_p20; 1.

```
DR Pfam; PF02758; PAAD DAPIN; 1.
DR PRINTS; PRO0376; IL1BCENZYME.
DR SMART; SW00115; CASC; 1.
DR PROSITE; PS01122; CASPASE_CYS; 1.
DR PROSITE; PS01121; CASPASE_HIS; 1.
DR PROSITE; PS0207; CASPASE_P10; 1.
DR PROSITE; PS0208; CASPASE_P20; 1.
DR PROSITE; PS0824; DAPIN; 1.
DR SEQUENCE 404 AA; 46077 MW; 629BECFFB4A87D3B CRC64;

Query Match      10.5%; Score 104; DB 13; Length 404;
Best Local Similarity 30.8%; Pred. No. 0.38;
Matches 36; Conservative 17; Mismatches 50; Indels 14; Gaps 4;

QY 10 DALENTARELKKFKLKLISVPLREGYGRIPRGALLSMDALDITDKLVSFYLETYGAELT 69
DB 10 DVLEDDVEAELEKQF-TROLWIGVKVGEPIPRGKLENKORQDVVDSMWQVQSEDAQT-IT 67
QY 70 ANVLRDMGLQEMAGQL-----QAATHQSGAAPAGIQAPPQSAAKPGLHFIHQ 117
DB 68 VQTLRKIKQNERAKRLNESLLKVSQGENKQNSEEP---QPIPIIISQPIQIIISQ 121

RESULT 8
Q8VPY3 PRELIMINARY; PRT; 127 AA.
AC Q8VPY3;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE SILE (Fragment).
GN SILE.
OS Serratia marcescens.
OG Plasmid pR478.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Serratia.
OX NCBI_TaxID=615;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=21602765; PubMed=11739772;
RA Gupta A., Phung L.T., Taylor D.E., Silver S.;
RT "Diversity of silver resistance genes in Inch incompatibility group plasmids.";
RL Microbiology 147:3393-3402(2001).
DR EMBL; AY009377; AAL68934.1; -.
KW Plasmid.
FT NON TER 127
SQ SEQUENCE 127 AA; 13438 MW; 67B3822C25BFC9A5 CRC64;

Query Match      9.2%; Score 91.5; DB 2; Length 127;
Best Local Similarity 31.6%; Pred. No. 1.1;
Matches 31; Conservative 11; Mismatches 47; Indels 9; Gaps 4;

QY 64 YGAELTANVLRDMGLQEMAGQLQAATHQ-GSGAAPAGIQAPPQSAAKPGLHFIHQRAAL 122
DB 12 FGLISSAWATETVNIHERVNNAPAHQMSAAAPVGIO-----GTAPRMAGMDQHEQAI 66

QY 123 IARVTNVEWLLDALYGVLTDEOYQ--AVRAEPTNPSK 158
DB 67 IAHTMTNGSADA-HQKMWESHQRMGSGQTSVPTGFSK 103

RESULT 9
Q8VPW8 PRELIMINARY; PRT; 127 AA.
AC Q8VPW8;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE SILE (Fragment).
GN SILE.
OS Salmonella oranienberg.
OG Plasmid MIP235.

QY 64 YGAELTANVLRDMGLQEMAGQLQAATHQ-GSGAAPAGIQAPPQSAAKPGLHFIHQRAAL 122
DB 12 FGLISSAWATETVNIHERVNNAPAHQMSAAAPVGIO-----GTAPRMAGMDQHEQAI 66

QY 123 IARVTNVEWLLDALYGVLTDEOYQ--AVRAEPTNPSK 158
DB 67 IAHTMTNGSADA-HQKMWESHQRMGSGQTSVPTGFSK 103

RESULT 10
Q8VPX8 PRELIMINARY; PRT; 127 AA.
AC Q8VPX8;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE SILE (Fragment).
GN SILE.
OS Salmonella enterica subsp. enterica serovar Ohio.
OG Plasmid pMIP233.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=117541;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=21602765; PubMed=11739772;
RA Gupta A., Phung L.T., Taylor D.E., Silver S.;
RT "Diversity of silver resistance genes in Inch incompatibility group plasmids.";
RL Microbiology 147:3393-3402(2001).
DR EMBL; AY009382; AAL68941.1; -.
KW Plasmid.
FT NON TER 127
SQ SEQUENCE 127 AA; 13456 MW; 67B3822C2313A365 CRC64;

Query Match      9.2%; Score 91.5; DB 2; Length 127;
Best Local Similarity 31.6%; Pred. No. 1.1;
Matches 31; Conservative 11; Mismatches 47; Indels 9; Gaps 4;

QY 64 YGAELTANVLRDMGLQEMAGQLQAATHQ-GSGAAPAGIQAPPQSAAKPGLHFIHQRAAL 122
DB 12 FGLISSAWATETVNIHERVNNAPAHQMSAAAPVGIO-----GTAPRMAGMDQHEQAI 66

QY 123 IARVTNVEWLLDALYGVLTDEOYQ--AVRAEPTNPSK 158
DB 67 IAHTMTNGSADA-HQKMWESHQRMGSGQTSVPTGFSK 103

RESULT 11
Q8VPY8 PRELIMINARY; PRT; 127 AA.
AC Q8VPY8;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DE SILE (Fragment).
GN SILE.
OS Salmonella enterica subsp. enterica serovar Ohio.
OG Plasmid pMIP233.

QY 64 YGAELTANVLRDMGLQEMAGQLQAATHQ-GSGAAPAGIQAPPQSAAKPGLHFIHQRAAL 122
DB 12 FGLISSAWATETVNIHERVNNAPAHQMSAAAPVGIO-----GTAPRMAGMDQHEQAI 66

QY 123 IARVTNVEWLLDALYGVLTDEOYQ--AVRAEPTNPSK 158
DB 67 IAHTMTNGSADA-HQKMWESHQRMGSGQTSVPTGFSK 103
```

DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
 DE SILE (Fragment).
 GN SILE.
 OS Serratia marcescens.
 OG Plasmid pR476b.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Serratia.
 OX NCBI_TaxID=615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21602765; PubMed=11739772;
 RA Gupta A., Phung L.T., Taylor D.E., Silver S.;
 RT "Diversity of silver resistance genes in *inch* incompatibility group
 RL plasmids";
 DR EMBL; AY009372; AAL68931.1; --
 KW Plasmid.
 FT NON_TER.
 SQ SEQUENCE 127 AA; 13456 MW; 67B3822C2313A365 CRC64;
 Query Match 9.2%; Score 91.5; DB 2; Length 127;
 Best Local Similarity 31.6%; Pred. No. 1.1;
 Matches 31; Conservative 11; Mismatches 47; Indels 9; Gaps 4;
 QY 64 YGAELTANVLRDMGLQEMAGQIQAATHQ-GSGAAPAGIQAPPSAAKPGHFIHQHRAAL 122
 DB 12 FGLISSAWATETVNIHERVNNVQAQAHOQMSAAAPVGIQ-----GTAPRMAGMDQHEQAI 66
 QY 123 IARVTNVEWLLDALYGVKVLTDQYQ--AVRAEPTNPSK 158
 DB 67 IAHETMTNGSADA-HQKMWESHQRMMGMSQTVSPTGFSK 103
 RESULT 12
 QY19L7 PRELIMINARY; PRT; 383 AA.
 ID Q919L7;
 AC Q919L7;
 DT 01-OCT-2000 (TremBLrel. 15, Created)
 DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
 DE Caspase.
 GN CASPA OR CASPY.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 CC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20373792; PubMed=10917738;
 RA Inohara N., Nunez G.;
 RT "Genes with homology to mammalian apoptosis regulators identified in
 RL zebrafish";
 RT Cell Death Differ. 7:509-510(2000).
 DR EMBL; AF233434; AAF66964.1; --
 DR HSP; P29466; 1ICE.
 DR ZFIN; ZDB-GENE-000616-3; caspa.
 DR InterPro; IPR002398; ICE.
 DR InterPro; IPR002138; ICE_P10.
 DR InterPro; IPR001309; ICE_P20.
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR Pfam; PF00655; ICE_P10; 1.
 DR Pfam; PF00856; ICE_P20; 1.
 DR Pfam; PF02758; PAAD_DAPIN; 1.
 DR PRINTS; PR00376; ILIBCENZYME.
 DR SMART; SM00115; CASC; 1.
 DR PROSITE; PS01122; CASPASE_CYS; 1.
 DR PROSITE; PS01121; CASPASE_HIS; 1.
 DR PROSITE; PS0207; CASPASE_P10; 1.
 DR PROSITE; PS0208; CASPASE_P20; 1.
 DR PROSITE; PS0824; DAPIN; 1.
 SQ SEQUENCE 383 AA; 43966 MW; 21890871309774C3 CRC64;

Query Match 9.2%; Score 91.5; DB 13; Length 383;
 Best Local Similarity 26.2%; Pred. No. 5;
 Matches 28; Conservative 24; Mismatches 44; Indels 11; Gaps 4;
 QY 5 RDAILDALENLTAEELKKFKLLSVPLREGVGRIPRGALLSM-DALDLTDKLVSFYLET 63
 DB 6 KHLQDALNSIGADNLRFPQRLGD---RKQEPVRKSTIEKLKDEIDLVDLVNFTSD 62
 QY 64 YGAELTANVLRDMGLQEMAGQIQAATHQSGAAPAGIQAPPSAAK 110
 DB 63 -AVSVTVDIRGIKCNAVAEELLENTGOG-----GVSQPEPPVPEP 102
 RESULT 13
 P74033 PRELIMINARY; PRT; 439 AA.
 ID P74033;
 AC P74033;
 DT 01-FEB-1997 (TremBLrel. 02, Created)
 DT 01-FEB-1997 (TremBLrel. 02, Last sequence update)
 DT 01-OCT-2002 (TremBLrel. 22, Last annotation update)
 DE Hypothetical protein slr0806.
 GN SLR0806.
 OS Synechocystis sp. (strain PCC 6803).
 CC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hirose M., Sugita M., Saito S., Kimura T.,
 RA Hosouchi T., Matsumoto A., Nakazaki N., Naruo K., Okumura S.,
 RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 RA Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 DR EMBL; D90911; BAA18106.1; --
 DR InterPro; IPR006094; Oxid_FAD_bind.
 DR Pfam; PF01565; FAD binding 4; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 439 AA; 47989 MW; 6C2E6F56365FA653 CRC64;
 Query Match 8.7%; Score 86.5; DB 16; Length 439;
 Best Local Similarity 24.0%; Pred. No. 17;
 Matches 42; Conservative 22; Mismatches 60; Indels 51; Gaps 8;
 QY 31 PLREGYGR---IPRGALLSMDALDITDKLVSYLETYGAEELTANVLRDMGLQEMAGQLQA 87
 DB 194 PVHGYGTNGIITEITLPLPALPWAEIVSF-----TNLSAIAFAQN 237
 QY 88 ATHQ-GSGAAPAGIQAPP-----QSNAPGLHFIHQHRAALIAVTNVEW-----LL 133
 DB 238 LAHQDGIYSKEISIQADPIQVFSLSKYQPGAHW-----MVIVSELDWLAFTQLA 290
 QY 134 DALYGVKVLTDQYQAVRAEPTNPSKMKLFSTPAWNWTKDLLLQALRESOSYL 188
 DB 291 KASKGEIIFEQ-----DPQSPGKKINLIEP-----NWNHTLLARAVDPSLTYL 334
 RESULT 14
 Q8NQ6 PRELIMINARY; PRT; 828 AA.
 ID Q8NQ6;
 AC Q8NQ6;
 DT 01-OCT-2002 (TremBLrel. 22, Created)
 DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
 DE Phenylalanyl-tRNA synthetase beta subunit (EC 6.1.1.20).
 GN CGL1390.
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 CC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=1718;


```

RN  [1]
RC  STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA  Nakagawa S.;
RT  "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL  Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF005278; BAB98783.1; -
DR  InterPro; IPR005146; B3 4.
DR  InterPro; IPR005147; B5-
DR  InterPro; IPR005121; Fdx-AntiCB.
DR  InterPro; IPR004532; Phet_Bact.
DR  InterPro; IPR002547; tRNA_bind.
DR  Pfam; PF03483; B3 4; 1.
DR  Pfam; PF03484; B5; 1.
DR  Pfam; PF03147; FDX-ACB; 1.
DR  Pfam; PF01588; tRNA_bind; 1.
DR  TIGRFAMs; TIGR00472; phet_bact; 2.
KW  Aminoacyl-tRNA synthetase; Ligase; Complete proteome.
SQ  SEQUENCE 828 AA; 88611 MW; E3F13B39CE5F29FE CRC64;

Query Match      8.7%; Score 86.5; DB 16; Length 828;
Best Local Similarity 24.5%; Pred.No. 41;
Matches 40; Conservative 29; Mismatches 73; Indels 21; Gaps 7;

QY  29  SVPLREGYGRIPRGALLSMDALDLDTKLVSYFL---ETYGAEILTANVLDMGLQEMAGOL 85
DB  : : ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
107 AISARETYGRMGAGMICSASELGLADKQNSGIIITLDPYSG-EPGEDARQALGLEDTVPDV 165
QY  86  QAATHQSGGAAPAGIQAPQSAAKPGLHFDIDHRAALIAARVTNVVLLDALYGVKLT--- 142
DB  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
166 NVTDPGRGALSARGLTRELASAF--SLTFTD---PAIEPAVAGIEVKVPVAVEGSLINVEL 220
QY  143  DEQYQAVRAEPTNPSPKMKLFSTPAWN---WTCKDLLQLALR 182
DB  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
221 REETRAIR-----FGLRKVSGIDPAESPFWMQRELMLSGQR 257

RESULT 15
Q8TSR7 PRELIMINARY; PRT; 580 AA.
AC  Q8TSR7;
DT  01-JUN-2002 (TrEMBLrel. 21, Created)
DT  01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT  01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE  DNA ligase (ATP).
GN  LiG OR MA0728.
OS  Methanosarcina acetivorans.
OC  Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC  Methanosarcinaceae; Methanosarcina.
OX  NCBI_TaxID=2214;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=C2A / ATCC 35395 / DSM 2834;
RX  MEDLINE=21929760; PubMed=11932238;
RA  Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA  FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA  Allen N., Naylor J., Stange-Thomann N., DeArellano K., Johnson R.,
RA  Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA  Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA  Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
RA  Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA  Springer T.A., Umavam L.A., White O., White R.H., de Macario E.C.,
RA  Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA  Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA  Metcalf W.W., Birren B.;
RT  "The genome of Methanosarcina acetivorans reveals extensive metabolic
RL  and physiological diversity.";
RL  Genome Res. 12:532-542(2002).
DR  EMBL; AE010734; AAM04168.1; -
DR  InterPro; IPR000977; DNA_ligase.
DR  InterPro; IPR002016; Peroxidase.
DR  Pfam; PF01068; DNA_ligase; 1.
DR  Pfam; PF04679; DNA_ligase_A_C; 1.

DR  Pfam; PF04675; DNA_ligase_A_N; 1.
DR  TIGRFAMs; TIGR00574; dna11; 1.
DR  PROSITE; PS00697; DNA_LIGASE_A1; 1.
DR  PROSITE; PS00160; DNA_LIGASE_A3; 1.
DR  PROSITE; PS00435; PEROXIDASE_1; 1.
KW  Ligase; Complete proteome.
SQ  SEQUENCE 580 AA; 65351 MW; 31C511F878C79E73 CRC64;

Query Match      8.7%; Score 86; DB 17; Length 580;
Best Local Similarity 23.1%; Pred.No. 28;
Matches 45; Conservative 30; Mismatches 60; Indels 50; Gaps 9;

QY  10  DALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLDTKLVSYFLETYGAELT 69
DB  : : ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
136 DILQKATPEE-GKVIIRIVLGRRLGFGD-----QFLLEAFSIAFT 175
QY  70  A-----NVLDMGLQEMAGLOAATHQSGGAAPAGIQAPQSAAKPGLHFDIQ 117
DB  : : ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
176 GDKKHAGKIKESYSVCTDIG--ELA---QTLAEHGAG-----APGYFSIKPGRPV--- 220
QY  118  HRAALIAARVTNVVLLDALYGVKLTDEQYQAVRAEPTNPSPKMKLFSTPAWNMTCKDL- 176
DB  : : ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
221 -KSMIAQRVESFEELEERIKGKKAEEKYDGERVQ-----IHKAGDEIKAFSRLEIDIT 273
QY  177  -----LLQALRESQS 186
DB  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
274 AQYPDIIIEAVRESIS 288

Search completed: January 29, 2004, 13:48:13
Job time : 12.8067 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 13:11:31 ; Search time 3.84236 Seconds
(without alignments)
2147.276 Million cell updates/sec

Title: US-09-996-617-8

Perfect score: 990

Sequence: 1 MGRARDAILDALLENLTAEEL.....LLQALRESQSYLVEDLERS 195

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/prodata/1/aa/5A COMB.pdb.*
- 2: /cgn2_6/prodata/1/aa/5B COMB.pdb.*
- 3: /cgn2_6/prodata/1/aa/6A COMB.pdb.*
- 4: /cgn2_6/prodata/1/aa/6B COMB.pdb.*
- 5: /cgn2_6/prodata/1/aa/6C COMB.pdb.*
- 6: /cgn2_6/prodata/1/aa/6D COMB.pdb.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	990	100.0	195	US-09-340-620A-49	Sequence 49, Appl
2	684	69.1	193	US-09-340-620A-61	Sequence 61, Appl
3	378	38.2	71	US-09-340-620A-58	Sequence 58, Appl
4	264	26.7	70	US-09-340-620A-57	Sequence 57, Appl
5	264	26.7	70	US-09-340-620A-66	Sequence 66, Appl
6	151.5	15.3	109	US-09-340-620A-71	Sequence 71, Appl
7	94.5	9.5	1233	US-09-328-352-7874	Sequence 7874, Ap
8	86.5	8.7	372	US-07-813-584A-3	Sequence 3, Appl
9	86.5	8.7	372	US-08-330-515-3	Sequence 3, Appl
10	84.5	8.5	284	US-09-069-023-5	Sequence 5, Appl
11	84.5	8.5	478	US-09-069-023-4	Sequence 4, Appl
12	84.5	8.5	530	US-09-069-023-3	Sequence 3, Appl
13	84.5	8.5	531	US-09-069-023-1	Sequence 1, Appl
14	84.5	8.5	540	US-09-019-942-1	Sequence 1, Appl
15	84.5	8.5	540	US-09-099-041A-2	Sequence 2, Appl
16	84.5	8.5	540	US-09-069-023-27	Sequence 27, Appl
17	84.5	8.5	540	US-09-245-281-2	Sequence 2, Appl
18	84.5	8.5	540	US-09-470-271-1	Sequence 1, Appl
19	84.5	8.5	540	US-09-207-359B-2	Sequence 2, Appl
20	84.5	8.5	540	US-09-340-620A-2	Sequence 2, Appl
21	84.5	8.5	540	US-09-345-473E-28	Sequence 28, Appl
22	78.5	7.9	424	US-09-107-532A-6238	Sequence 6238, Ap
23	77.5	7.8	2436	US-08-444-818-75	Sequence 75, Appl
24	76.5	7.7	164	US-09-245-281-41	Sequence 41, Appl
25	76.5	7.7	164	US-09-207-359B-41	Sequence 41, Appl
26	76.5	7.7	164	US-09-340-620A-41	Sequence 41, Appl
27	76.5	7.7	249	US-09-245-281-39	Sequence 39, Appl

28	76.5	7.7	249	4	US-09-207-359B-39	Sequence 39, Appl
29	76.5	7.7	249	4	US-09-340-620A-39	Sequence 39, Appl
30	76.5	7.7	409	4	US-09-207-359B-46	Sequence 46, Appl
31	76.5	7.7	953	4	US-09-099-041A-8	Sequence 8, Appl
32	76.5	7.7	953	4	US-09-245-281-8	Sequence 8, Appl
33	76.5	7.7	953	4	US-09-207-359B-8	Sequence 8, Appl
34	76.5	7.7	953	4	US-09-340-620A-8	Sequence 8, Appl
35	76	7.7	435	4	US-09-252-991A-31215	Sequence 31215, A
36	76	7.7	1175	4	US-09-069-023-6	Sequence 25044, A
37	75.5	7.6	167	4	US-09-069-023-6	Sequence 6, Appl
38	75.5	7.6	3011	3	US-08-811-566-20	Sequence 20, Appl
39	75.5	7.6	3011	3	US-09-014-416-1	Sequence 1, Appl
40	75.5	7.6	3011	3	US-09-014-416-5	Sequence 5, Appl
41	75.5	7.6	3011	4	US-09-034-756-20	Sequence 20, Appl
42	75.5	7.6	3012	3	US-08-811-566-2	Sequence 2, Appl
43	75.5	7.6	3012	4	US-09-034-756-2	Sequence 2, Appl
44	75	7.6	323	4	US-09-252-991A-32273	Sequence 32273, A
45	75	7.6	372	4	US-09-328-352-5345	Sequence 5345, Ap

ALIGNMENTS

RESULT 1

US-09-340-620A-49
; Sequence 49, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-340-620A-49

Query Match 100.0%; Score 990; DB 4; Length 195;
Best Local Similarity 100.0%; Pred. No. 1.8e-112;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MGRARDAILDALLENLTAEELKKFKLLSVLPUREGYGRIPRGALLSMDALDITDKLVSPY	60
Db	1	MGRARDAILDALLENLTAEELKKFKLLSVLPUREGYGRIPRGALLSMDALDITDKLVSPY	60
QY	61	LETYCAELTANVLRDMGLQEMAGQIQAAATHOGSGGAAPAGIQAPPOSAAPKGLHFTDOHRA	120
Db	61	LETYCAELTANVLRDMGLQEMAGQIQAAATHOGSGGAAPAGIQAPPOSAAPKGLHFTDOHRA	120
QY	121	ALIARTVNVEMLLDALYKGLVTDEQYQAVRAEPTNPSKMKLFSTPAWNTCKDLLQA	180
Db	121	ALIARTVNVEMLLDALYKGLVTDEQYQAVRAEPTNPSKMKLFSTPAWNTCKDLLQA	180
QY	181	LRESQSYLVEDLERS	195
Db	181	LRESQSYLVEDLERS	195

RESULT 2

US-09-340-620A-61
; Sequence 61, Application US/09340620A

```
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-340-620A-61

Query Match      69.1%; Score 684; DB 4; Length 193;
Best Local Similarity 71.8%; Pred. No. 3.5e-75;
Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;

QY 1 MGRARDAILDALENLTABELKKFKLLSVPLREGYGRIPRGALLSMDALDLTKLVSPY 60
DB 1 MGRARDAILDALENLTABELKKFKMKLLTVQLREGYGRIPRGALLQMDAIDLTKLVSY 60
QY 61 LETYGAEALTANVLDRMGLOEMAGLOAAHTOGSGAAPAGIQAPPOSAAKPGHLHFIDQHEA 120
DB 61 LESYGLELTWTVLRDMGLQELAEQLQ-TTKESGAVAAAASVPAQTATGTG-HFVDQHRQ 118
QY 121 ALIARTVNVWLLDALYKGLTDEQYQAVRAEPTNPSKMKLFSFTPAWNTCKDLLLLQA 180
DB 119 ALIARTVTEVDGLDALHGSVLTGEGYQAVRAETTSQDKMKLFSFVPSNNLTCKDLSLLQA 178
QY 181 LRESQSYLVEDLERS 195
DB 179 LKEIHPYLVMDLEQS 193

RESULT 3
US-09-340-620A-58
; Sequence 58, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 58
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-340-620A-58

Query Match      38.2%; Score 378; DB 4; Length 71;
Best Local Similarity 100.0%; Pred. No. 1.5e-38;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 111 GLHFIDQHRRAALIARTVNVWLLDALYKGLTDEQYQAVRAEPTNPSKMKLFSFTPAWN 170
DB 1 GLHFIDQHRRAALIARTVNVWLLDALYKGLTDEQYQAVRAEPTNPSKMKLFSFTPAWN 60
QY 171 WTKCDLLLOAL 181
DB 61 WTKCDLLLOAL 71
```

```
RESULT 4
US-09-340-620A-57
; Sequence 57, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 57
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-340-620A-57
```

```
Query Match      26.7%; Score 264; DB 4; Length 70;
Best Local Similarity 73.9%; Pred. No. 1.2e-24;
Matches 51; Conservative 7; Mismatches 11; Indels 0; Gaps 0;
```

```
QY 113 HFIDQHRRAALIARTVNVWLLDALYKGLTDEQYQAVRAEPTNPSKMKLFSFTPAWNT 172
DB 2 HFVDQHRQALIARTVTEVDGLDALHGSVLTGEGYQAVRAETTSQDKMKLFSFVPSNNLT 61
QY 173 CKDLLLLQAL 181
DB 62 CKDLSLLQAL 70
```

```
RESULT 5
US-09-340-620A-66
; Sequence 66, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 66
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Mus musculus
```

```

US-09-340-620A-66
Query Match          26.7%; Score 264; DB 4; Length 70;
Best Local Similarity 73.9%; Pred. No. 1.2e-24;
Matches 51; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 113 HFIDHRAALIAIRVTVNVEWLLDALYGVKVLTDQYQAVRAEPTNPSKMRKLFPSFTFAWNT 172
DB 2 HFVDHQALIAIRVTVEDGVLDALHGSVLTQGYQAVRAETTSQDKMRKLFPSFVSNLT 61
QY 173 CKDILLQAL 181
DB 62 CKDILLQAL 70

RESULT 6
US-09-340-620A-71
Sequence 71, Application US/09340620A
Patent No. 6482933
GENERAL INFORMATION:
APPLICANT: Bertin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
FILE REFERENCE: 07334-124001
CURRENT APPLICATION NUMBER: US/09/340,620A
CURRENT FILING DATE: 1999-06-28
PRIOR APPLICATION NUMBER: US 09/245,281
PRIOR FILING DATE: 1998-02-05
PRIOR APPLICATION NUMBER: US 09/207,359
PRIOR FILING DATE: 1998-12-08
PRIOR APPLICATION NUMBER: US 09/099,041
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: US 09/019,942
PRIOR FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 71
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 71
LENGTH: 109
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Consensus sequence
NAME/KEY: VARIANT
LOCATION: (1)...(109)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-340-620A-71

Query Match          15.3%; Score 151.5; DB 4; Length 109;
Best Local Similarity 46.7%; Pred. No. 1.2e-10;
Matches 42; Conservative 13; Mismatches 28; Indels 7; Gaps 4;

QY 111 GLHFIDHRAALIAIRVTN--VEWLLDALYGVK-VLTDEQYQAVRAEPTNPSKMRKLFPSFTF 167
DB 7 GSEIDQHRAALIAIRVTEPDXSLLDALLSRDLISEDYEAETTXLSKVRKLLILVQ 66
QY 168 A-WNWTCKDL---LLOALRESOSYLVEDLE 193
DB 67 SKGEETCKXFLKLLQALQKDAAYLGLDPE 96

RESULT 7
US-09-328-352-7874
Sequence 7874, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 9252
SEQ ID NO 7874
LENGTH: 1233

```

```

; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-7874

Query Match          9.5%; Score 94.5; DB 4; Length 1233;
Best Local Similarity 26.5%; Pred. No. 0.047; 76; Indels 45; Gaps 10;
Matches 53; Conservative 26; Mismatches 26; Indels 45; Gaps 10;

QY 8 ILDALENITABELKKFKLLSVPLREGYGRIPRGALLSMDALDLTKLVSFYLETYGAE 67
DB 125 VIDQLDELFGVTLDSFSQKLLREFAFES-GKIER-AQITDDA-----KTYSRQ 170
QY 68 LTANVLRD-----MGLQEMAGLOQA-----THQSGAAPAGIQAPQSAKPG 112
DB 171 LTHDVLREWIQSQPTVIDALYLAGEKLSVDSFVKLVEDSLNFSAHFKLP-----EKPTI 226
QY 113 HFIDHRAALIAIRVTVNVEWLLDALYGVKVLTDQYQAVRAEPTNPSKMRKLFPSFTFAWNT 171
DB 227 QF--EQLAQLKQLATEIDISLEPY--LLOGEYKHVNGTIFRNGAFNKLFS----- 275
QY 172 TKCOLLLQALRESOSYLVED 191
DB 276 ECLPQLQLKQSDSILVFD 295

RESULT 8
US-07-813-584A-3
Sequence 3, Application US/07813584A
Patent No. 5352588
GENERAL INFORMATION:
APPLICANT: Fischetti, Vincent A.
APPLICANT: Bessen, Debra E.
TITLE OF INVENTION: No. 5352588e1 Immunoglobulin A Binding Protein
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESS: Kittle Murray
STREET: 98 Cutter Mill Road
CITY: Great Neck
STATE: NY
COUNTRY: USA
ZIP: 11021
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/07/813,584A
APPLICATION NUMBER: 19911224
FILING DATE: 19911224
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murray, Kittle
REGISTRATION NUMBER: 30,246
REFERENCE/DOCKET NUMBER: RU-100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-482-1990
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 372 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-813-584A-3

Query Match          8.7%; Score 86.5; DB 1; Length 372;
Best Local Similarity 26.8%; Pred. No. 0.068;
Matches 53; Conservative 30; Mismatches 78; Indels 37; Gaps 9;

QY 12 LENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDAL-----DLTKLVSFYLETYGA 66
DB 78 LEKINAEBEKKNKLEKAINELNENYKLDG-----IDALEKEKEDLTKTTAKTKNEIS 133
QY 67 ELTANVL-RDMGLQEMAGLOQAATHQSGAAPAGIQAPQ--SAAKPGL-HFIDHRAAL 122

```

Db 134 EASRGLSRDLEASRTAKKELEAKHQLEAKNKLTEGQVSEASRKGSLNDLEASRAA- 192
QY 123 IARVTNVEWLLDALYKGLVTDEQ-----YQAVRAEPTNPSKMRKLPSPFPANNWTKDOLL 177
Db 193 -----KKELEAKYQKLETDHQALEAKHQKLEADYQVSETSRKGLS-----RD-- 234
QY 178 LQALRESQSVLVEDLRS 195
Db 235 LEASREANKVTSelta 252

RESULT 9
US-08-330-515-3
; Sequence 3, Application US/08330515
; Patent No. 5556944
; GENERAL INFORMATION:
; APPLICANT: Fischetti, Vincent A.
; APPLICANT: Bessen, Debra E.
; TITLE OF INVENTION: No. 5556944el Immunoglobulin A Binding Protein
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/330,515
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,584
; FILING DATE: 24-DEC-1991
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: RU-100.D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 372 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-330-515-3

Query Match 8.7%; Score 86.5; DB 1; Length 372;
Best Local Similarity 26.8%; Pred. No. 0.068;
Matches 53; Conservative 30; Mismatches 78; Indels 37; Gaps 9;
QY 12 LENLTAEELKPKKLLSVLPREGYGRIPRGALLSMDAL-----DLTDKLVSFYLETYGA 66
Db 78 LEKINAEENKKLEAINKELNENYVKLDG-----IDALEKEKEDKTTLAKTTVENEIS 133
QY 67 ELTANYLV-RDMGLQEWAGLQATHOGSGAAPAGIQAPPQ--SAAKPGL-HFIDQHRAL 122
Db 134 EASRGLSRDLEASRTAKKELEAKHQLEAKNKLTEGQVSEASRKGSLNDLEASRAA- 192
QY 123 IARVTNVEWLLDALYKGLVTDEQ-----YQAVRAEPTNPSKMRKLPSPFPANNWTKDOLL 177
Db 193 -----KKELEAKYQKLETDHQALEAKHQKLEADYQVSETSRKGLS-----RD-- 234
QY 178 LQALRESQSVLVEDLRS 195

Db 235 LEASREANKVTSelta 252
RESULT 10
US-09-069-023-5
; Sequence 5, Application US/09069023A
; Patent No. 6348573
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; APPLICANT: Koseki, Takeyoshi
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
; FILE REFERENCE: UM-03333
; CURRENT APPLICATION NUMBER: US/09/069,023A
; CURRENT FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-069-023-5
Query Match 8.5%; Score 84.5; DB 4; Length 284;
Best Local Similarity 20.1%; Pred. No. 0.079;
Matches 50; Conservative 33; Mismatches 71; Indels 95; Gaps 9;
QY 6 DAILDALENLTAE-----LKKFKLLSVLPREGYGRIPRGALLSMDALDLDTKLVSPYL 61
Db 35 EPVLTTFEITFLEAVIQLKTKLQSVS-----SAIHLCDK----- 70
QY 62 ETYGAELTANVLRDMGLQ-EMAGLQATHOGSGAAPAGIOAP----- 103
Db 71 --KRMELSLNIPVNHGPOEESGSQLHENGSPETSLPAPQDNDFLSRKAQDCYPMK 128
QY 104 -----POSAA-----KPGI--HFIDQHR 119
Db 129 LHHCPGNHSDSTISSGQRAAFCDHKHTPCSSAINPLSTAGNERLQPGIAQWQSKR 188
QY 120 AALIARVTN--VEWLLDALYK-VLTDEQYQAVRAEPTNPSKMRKLPSPFPANNWTKDOL 176
Db 189 EDIVNQMTAEACLNQSLDALLSRDLIMKEDYELVSTKPTRTTSKVRQLDITDIQGEPAKV 248
QY 177 LQALRESQ 185
Db 249 IVQKLKONK 257
RESULT 11
US-09-069-023-4
; Sequence 4, Application US/09069023A
; Patent No. 6348573
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; APPLICANT: Koseki, Takeyoshi
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
; FILE REFERENCE: UM-03333
; CURRENT APPLICATION NUMBER: US/09/069,023A
; CURRENT FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-069-023-4
Query Match 8.5%; Score 84.5; DB 4; Length 478;
Best Local Similarity 20.1%; Pred. No. 0.18;

```
Matches 50; Conservative 33; Mismatches 71; Indels 95; Gaps 9
```

QY 6 DAIIDALENTABE-----LKKFKLLKLSVLPUREGYGRIPRGALLSMDALDITDKLVSPYL 61
 : : : : : : : : : : : : : : :
Db 229 EPVLRTFEEITFLFAVIQLKKTKLQSVS-----SAIHLCDK-----264
 : : : : : : : : : : : : : :
QY 62 ETYGAEILTANVLRDMGLQ-EMAGOLQAATHOGSGAAPAGIOAP-----103
 : : : : : : : : : : : : : :
Db 265 --KKMELSLNIPVNHPQPQESCGSQLHENGSPETSRSLPAPODNFLSRKAQDCYFMK 322
 : : : : : : : : : : : : : :
QY 104 -----PQSAA-----KPGL--HFIDQHR 119
 : : : : : : : : : : : : : :
Db 323 LHHCPGNHSWDSTTIGSQRAAFCDHKTTPTCSSAINPLSTAGNSERLOPGIAQQWIQSKR 382
 : : : : : : : : : : : : : :
QY 120 AALTARVTN--VENLLDALYCK-VLTDEYOQAVRAEPNPSMKRLFSFTPAWNWTCKOL 176
 :: :: : : : : : : : : : : : :
Db 383 EDIVNQMTAECLNOSLDALLSRDLIMKEDYELVSTKPTRTSKVRLDITTDIQGEFAKV 442
 : : : : : : : : : : : : : :
QY 177 LLQALRESQ 185
 : : : : : :
Db 443 IVQKLKONK 451

RESULT 12
US-09-069-023-3
; Sequence 3, Application US/09069023A
; Patent No. 6348573
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; APPLICANT: Koseki, Takeyoshi
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
; TITLE OF INVENTION: SIGNALING PATHWAY INHIBITORS AND ACTIVATORS
; FILE REFERENCE: UM-03333
; CURRENT APPLICATION NUMBER: US/09/069,023A
; CURRENT FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 530
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-069-023-3

	Query Match	8.5%; Score 84.5; DB 4; Length 530;	
	Best Local Similarity	20.1%; Pred. No. 0.21;	
	Matches	50; Conservative 33; Mismatches 71; Indels 95; Gaps 9;	
Qy	6	DAILDALENLTAE-----LKKFKLLSVPLREGYGRIPRGALLSDALDLTDKLVSYFL	61
Db	281	EPVLRTFBEITFLEAVIQLKTKLQSVS-----SAIHLCDK-----	316
Qy	62	ETYGAELTANVLRDMGLQ-EMAGQLQAATHOGSSGAAPAGIOAP-----	103
Db	317	--KMELSINPVNHHGPOEESGSSLHENGSPETSRSLPAPQDNDFLSRKAQCYPMK	374
Qy	104	-----PQSAA-----KPGL--HFIDQHR	119
Db	375	LHHCPSHNHSWSTISGSQRAAFCDHKHTTPCSSAIINPLSTAGNSERLOPGTAQQWIQSKR	434
Qy	120	AALIARVTN--VEWLDDALYGK-VLTDEQYQAVRAEPTNPQMRKLFGFTPAWNWTCKDL	176
Db	435	EDIVNQMTACLNQSLDALLSRDLIMKEDIYELVSTKPTRTSKVRQLLDTDIQGEEFAKV	494
Qy	177	LLOALRESQ	185
Db	495	IIVQKLKDKNK	503

RESULT 13
US-09-069-023-1
; Sequence 1, Application US/09069023A
; Patent No. 6348573

```

: GENERAL INFORMATION:
: APPLICANT: Nunez, Gabriel
: APPLICANT: Inohara, Naohiro
: APPLICANT: Koseki, Takeyoshi
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
: TITLE OF INVENTION: SIGNALING PATHWAY INHIBITORS AND ACTIVATORS
: FILE REFERENCE: UM-0333
: CURRENT APPLICATION NUMBER: US/09/069,023A
: CURRENT FILING DATE: 1998-04-27
: NUMBER OF SEQ ID NOS: 38
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 531
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-069-023-1

Query Match      8.5%; Score 84.5; DB 4; Length 531;
Best Local Similarity 20.1%; Pred No.0.21; Indels 95; Gaps 9
Matches 50; Conservative 33; Mismatches 71;

Qy 6 DAILDALLENLTAAEE---LKKFKLLKLSVPLREGYGRIPRGALLSMDALDLTKLVSYL 61
Db 282 EPVLTREEITFLEAVIQLKTKLQSVS-----SAHLCDK----- 317

Qy 62 ETYGAELTANVLRDMGLQ-EMAGQLQAATHOGSGNAPAGIQAP----- 103
Db 318 --KMWELSLNIPVNHGPOEESCGSQLHENSGBPETSRLPAPQNDFLSRKAQDCYFMK 375

Qy 104 -----POSAA-----KPEGL--HFIDQHR 119
Db 376 LHHCPGNHWSDTTSGSORAFCDHKTTPCSAIIINLSTAGNSERLQPGIAQQWISKR 435

Qy 120 AALIARTVTN--VEWLLDALYKG-VLTDRQYQAVRAEPTNPSSKMKLFSFTPAWNTCKDL 176
Db 436 EDIVNQWTEACLNQSLDALLSRDLIMKEDYELVSTKPTRTSKVRQLDITDTCGEEFAKV 495

Qy 177 LQALRESQ 185
Db 496 IVQKLKDKN 504

RESULT 14
US-09-019-942-1
: Sequence 1, Application US/09019942
: Patent No. 6033855
: GENERAL INFORMATION:
: APPLICANT: Bertin, John
: TITLE OF INVENTION: GENES ENCODING CASPASE RECRUITMENT
: TITLE OF INVENTION: DOMAIN POLYPEPTIDES
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02110-2804
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: Windows 95
: SOFTWARE: FastSeq for Windows Version 2.0b
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/019,942
: FILING DATE: 06-FEB-1998
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Meiklejohn, Ph.D., Anita L.
: REGISTRATION NUMBER: 35,283
: REFERENCE/DOCKET NUMBER: 07334/068001

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 13:46:01 ; Search time 33.9809 Seconds
(without alignments)
1192.602 Million cell updates/sec

Title: US-09-996-617-8
Perfect score: 990
Sequence: 1 MGRARDAILDALENLTABEEL.....LLQALRESQSYLVEDLRS 195

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PTCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	990	100.0	195	9	US-09-728-721-49
2	990	100.0	195	10	US-09-996-617-8
3	990	100.0	195	10	US-09-841-879B-5
4	990	100.0	195	12	US-10-131-410-84
5	990	100.0	195	12	US-10-240-145-55
6	990	100.0	195	15	US-10-295-981-49
7	990	100.0	205	15	US-10-106-698-5421
8	759	76.7	190	9	US-09-925-301-1120
9	684	69.1	193	9	US-09-728-721-61
10	684	69.1	193	10	US-09-841-879B-2
11	684	69.1	193	15	US-10-295-981-61
12	564.5	57.0	158	15	US-10-106-698-5422
13	509	51.4	136	9	US-09-925-299-1421
14	509	51.4	136	11	US-09-925-299-1421
15	469	47.4	90	10	US-09-931-071-7

16	445	44.9	85	10	US-09-841-879B-8	Sequence 8, Appli
17	435	43.9	89	14	US-10-127-516-16	Sequence 16, Appl
18	435	43.9	89	14	US-10-027-629-16	Sequence 16, Appl
19	378	38.2	71	9	US-09-728-721-58	Sequence 58, Appl
20	378	38.2	71	15	US-10-295-981-58	Sequence 58, Appl
21	377	38.1	77	14	US-10-127-516-8	Sequence 8, Appli
22	377	38.1	77	14	US-10-027-629-8	Sequence 8, Appli
23	370	37.4	77	12	US-10-132-967-8	Sequence 8, Appli
24	368	37.2	76	11	US-09-965-621-9	Sequence 9, Appli
25	368	37.2	76	12	US-10-407-866-9	Sequence 9, Appli
26	309	31.2	85	10	US-09-841-879B-15	Sequence 15, Appl
27	301	30.4	84	10	US-09-841-879B-7	Sequence 7, Appli
28	267	27.0	89	11	US-09-965-621-28	Sequence 28, Appl
29	267	27.0	89	12	US-10-407-866-28	Sequence 28, Appl
30	267	27.0	89	14	US-10-127-516-17	Sequence 17, Appl
31	267	27.0	89	14	US-10-027-629-17	Sequence 17, Appl
32	264	26.7	70	9	US-09-728-721-57	Sequence 57, Appl
33	264	26.7	70	9	US-09-728-721-66	Sequence 66, Appl
34	264	26.7	70	15	US-10-295-981-57	Sequence 57, Appl
35	264	26.7	70	15	US-10-295-981-66	Sequence 66, Appl
36	250.5	25.3	1399	10	US-09-388-221-4	Sequence 4, Appli
37	250.5	25.3	1429	10	US-09-996-617-2	Sequence 2, Appli
38	250.5	25.3	1429	10	US-09-931-071-2	Sequence 2, Appli
39	250.5	25.3	1429	12	US-10-028-374-15	Sequence 15, Appl
40	250.5	25.3	1429	12	US-10-183-770-15	Sequence 15, Appl
41	250.5	25.3	1429	15	US-10-028-392-11	Sequence 11, Appl
42	250	25.3	1443	10	US-09-388-221-6	Sequence 6, Appli
43	250	25.3	1473	10	US-09-388-221-2	Sequence 2, Appli
44	249.5	25.2	1429	12	US-10-028-374-3	Sequence 3, Appli
45	249.5	25.2	1429	12	US-10-183-770-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-728-721-49
; Sequence 49, Application US/09728721
; Patent No. US20020061845A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/728,721
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-728-721-49

Query Match 100.0%; Score 990; DB 9; Length 195;
Best Local Similarity 100.0%; Pred. No. 3.1e-98;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGRARDAILDALENLTABEELKFKLLSVLPREGVGRIPRGALLSMDALDTDKLVSY 60
Db	1	MGRARDAILDALENLTABEELKFKLLSVLPREGVGRIPRGALLSMDALDTDKLVSY 60
Qy	61	LETYGAELTANVLRDMGLQEMAGQLQAATHQSGGAAPAGIQAPQSAAKPGHFFIDQHRA 120
Db	61	LETYGAELTANVLRDMGLQEMAGQLQAATHQSGGAAPAGIQAPQSAAKPGHFFIDQHRA 120

QY 121 ALIARVTNVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSTTPANNWTKDLLLLQA 180
DB 121 ALIARVTNVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSTTPANNWTKDLLLLQA 180
QY 181 LRESQSYLVEDLERS 195
DB 181 LRESQSYLVEDLERS 195

RESULT 2

US-09-996-617-8
; Sequence 8, Application US/09996617
; Patent No. US20020128198A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; FILE REFERENCE: 07334-340001
; CURRENT APPLICATION NUMBER: US/09/996,617
; CURRENT FILING DATE: 2001-11-27
; PRIOR APPLICATION NUMBER: 09/931,071
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/428,252
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-996-617-8

Query Match 100.0%; Score 990; DB 10; Length 195;
Best Local Similarity 100.0%; Pred. No. 3.1e-98;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRARDAILDALENLTABELKKFKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
DB 1 MGRARDAILDALENLTABELKKFKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
QY 61 LETYGAELTANVLRDMGLQEMAGOLQAATHOGSGAAGPAGIQAPPQSAAPKPGHFIHQHRA 120
DB 61 LETYGAELTANVLRDMGLQEMAGOLQAATHOGSGAAGPAGIQAPPQSAAPKPGHFIHQHRA 120
QY 121 ALIARVTNVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSTTPANNWTKDLLLLQA 180
DB 121 ALIARVTNVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSTTPANNWTKDLLLLQA 180
QY 181 LRESQSYLVEDLERS 195
DB 181 LRESQSYLVEDLERS 195

RESULT 3

US-09-841-8798-5
; Sequence 5, Application US/09841879B
; Patent No. US20020142979A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-330001
; CURRENT APPLICATION NUMBER: US/09/841,879B
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: US 09/728,721
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 195

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-841-879B-5

Query Match 100.0%; Score 990; DB 10; Length 195;
Best Local Similarity 100.0%; Pred. No. 3.1e-98;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRARDAILDALENLTABELKKFKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
DB 1 MGRARDAILDALENLTABELKKFKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
QY 61 LETYGAELTANVLRDMGLQEMAGOLQAATHOGSGAAGPAGIQAPPQSAAPKPGHFIHQHRA 120
DB 61 LETYGAELTANVLRDMGLQEMAGOLQAATHOGSGAAGPAGIQAPPQSAAPKPGHFIHQHRA 120
QY 121 ALIARVTNVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSTTPANNWTKDLLLLQA 180
DB 121 ALIARVTNVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSTTPANNWTKDLLLLQA 180
QY 181 LRESQSYLVEDLERS 195
DB 181 LRESQSYLVEDLERS 195

RESULT 4

US-10-131-410-84
; Sequence 84, Application US/10131410
; Publication No. US20030235915A1
; GENERAL INFORMATION:
; APPLICANT: SPECHT, THOMAS
; APPLICANT: HINZMANN, BERNI
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILARSKI, CHRISTIAN
; APPLICANT: DAHL, EDGAR
; APPLICANT: ROSENTHAL, ANDRE
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM TISSUE OF BREAST
; TITLE OF INVENTION: TUMORS
; FILE REFERENCE: SCH-1763
; CURRENT APPLICATION NUMBER: US/10/131,410
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 09/646,673
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: PCT/DE99/00908
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 202
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 84
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-410-84

Query Match 100.0%; Score 990; DB 12; Length 195;
Best Local Similarity 100.0%; Pred. No. 3.1e-98;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGRARDAILDALENLTABELKKFKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
DB 1 MGRARDAILDALENLTABELKKFKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
QY 61 LETYGAELTANVLRDMGLQEMAGOLQAATHOGSGAAGPAGIQAPPQSAAPKPGHFIHQHRA 120
DB 61 LETYGAELTANVLRDMGLQEMAGOLQAATHOGSGAAGPAGIQAPPQSAAPKPGHFIHQHRA 120
QY 121 ALIARVTNVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSTTPANNWTKDLLLLQA 180
DB 121 ALIARVTNVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSTTPANNWTKDLLLLQA 180
QY 181 LRESQSYLVEDLERS 195
DB 181 LRESQSYLVEDLERS 195

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-981-49

Query Match      100.0%; Score 990; DB 15; Length 195;
Best Local Similarity 100.0%; Pred. No. 3.1e-98;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MGRARDAILDALLENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
Db      1 MGRARDAILDALLENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
Qy      61 LETYGAELTANVLRDMGLQEMAGOLQAATHOGSGAAGPAGIQAPPOSAAKPGLHFIHQHRA 120
Db      61 LETYGAELTANVLRDMGLQEMAGOLQAATHOGSGAAGPAGIQAPPOSAAKPGLHFIHQHRA 120
Qy      121 ALIARVTNVVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSTPAWNWTCKDILLQA 180
Db      121 ALIARVTNVVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSTPAWNWTCKDILLQA 180
Qy      181 LRESQSYLVEDLERS 195
Db      181 LRESQSYLVEDLERS 195

RESULT 7
US-10-106-698-5421
; Sequence 5421, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 5421
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-5421

Query Match      100.0%; Score 990; DB 15; Length 205;
Best Local Similarity 100.0%; Pred. No. 3.4e-98;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MGRARDAILDALLENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
Db      11 MGRARDAILDALLENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 70
Qy      61 LETYGAELTANVLRDMGLQEMAGOLQAATHOGSGAAGPAGIQAPPOSAAKPGLHFIHQHRA 120
Db      71 LETYGAELTANVLRDMGLQEMAGOLQAATHOGSGAAGPAGIQAPPOSAAKPGLHFIHQHRA 130
Qy      121 ALIARVTNVVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSTPAWNWTCKDILLQA 180
Db      131 ALIARVTNVVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSTPAWNWTCKDILLQA 190
Qy      181 LRESQSYLVEDLERS 195
Db      191 LRESQSYLVEDLERS 205

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-981-49

Query Match      100.0%; Score 990; DB 15; Length 195;
Best Local Similarity 100.0%; Pred. No. 3.1e-98;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MGRARDAILDALLENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
Db      1 MGRARDAILDALLENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
Qy      61 LETYGAELTANVLRDMGLQEMAGOLQAATHOGSGAAGPAGIQAPPOSAAKPGLHFIHQHRA 120
Db      61 LETYGAELTANVLRDMGLQEMAGOLQAATHOGSGAAGPAGIQAPPOSAAKPGLHFIHQHRA 120
Qy      121 ALIARVTNVVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSTPAWNWTCKDILLQA 180
Db      121 ALIARVTNVVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSTPAWNWTCKDILLQA 180
Qy      181 LRESQSYLVEDLERS 195
Db      181 LRESQSYLVEDLERS 195

RESULT 6
US-10-295-981-49
; Sequence 49, Application US/10295981
; Publication No. US20030120055A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/10/295,981
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US/09/340,620
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-240-145-55

Query Match      100.0%; Score 990; DB 12; Length 195;
Best Local Similarity 100.0%; Pred. No. 3.1e-98;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MGRARDAILDALLENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
Db      1 MGRARDAILDALLENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLTDKLVSFY 60
Qy      61 LETYGAELTANVLRDMGLQEMAGOLQAATHOGSGAAGPAGIQAPPOSAAKPGLHFIHQHRA 120
Db      61 LETYGAELTANVLRDMGLQEMAGOLQAATHOGSGAAGPAGIQAPPOSAAKPGLHFIHQHRA 120
Qy      121 ALIARVTNVVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSTPAWNWTCKDILLQA 180
Db      121 ALIARVTNVVWLLDALYGVLTDEQYQAVRAEPTNPSKMRKLFSTPAWNWTCKDILLQA 180
Qy      181 LRESQSYLVEDLERS 195
Db      181 LRESQSYLVEDLERS 195

RESULT 5
US-10-240-145-55
; Sequence 55, Application US/10240145
; Publication No. US20030235883A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-048
; CURRENT APPLICATION NUMBER: US/10/240,145
; CURRENT FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: US/09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US/09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US/09/668,680
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/09/695,618
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: US/09/728,711
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: NOT YET ASSIGNED
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: Custom
; SEQ ID NO 55
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-240-145-55
```

RESULT 8

US-09-925-301-1120
; Sequence 1120, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1120
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1120

Query Match 76.7%; Score 759; DB 9; Length 190;
Best Local Similarity 95.6%; Pred. No. 2.1e-73;
Matches 153; Conservative 3; Mismatches 2; Indels 2; Gaps 1;

QY 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60
DB 11 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 70
QY 61 LETYGAEALTANVLRDMGLQEMAGOLQAATHOGSGAAPAGIQAPPOSAAKPGHLHFIDQHRA 120
DB 71 LETYGAEALTANVLRDMGLQEMAGOLQAATHOGSGAAPAGIQAPPOSAAKPGHLHFIDQHRA 130
QY 121 ALIARVTNVWLLDALYKGKVLTDQYQAVRAEPTNPSPKMR 160
DB 131 ALIARVTNVWLLDALYKGKVLTDQYQAVR--PSPPTQAR 168

RESULT 9

US-09-728-721-61
; Sequence 61, Application US/09728721
; Patent No. US20020061845A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/728,721
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-728-721-61

Query Match 69.1%; Score 684; DB 9; Length 193;
Best Local Similarity 71.8%; Pred. No. 2.6e-65;
Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;

QY 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60
DB 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLQMDAIDLTDKLVSY 60

QY 61 LETYGAEALTANVLRDMGLQEMAGOLQAATHOGSGAAPAGIQAPPOSAAKPGHLHFIDQHRA 120
DB 61 LETYGAEALTANVLRDMGLQELAEQLQ-TTKEESGAVAAAAASVPAQSTARTG-HFVDQHRQ 118
QY 121 ALIARVTNVWLLDALYKGKVLTDQYQAVRAEPTNPSPKMRKLFSTPAWNWTKDILLQA 180
DB 119 ALIARVTEVDGVLDAHGSVLTEGQYQAVRAETTSQDKMRKLFSPVPSWNLTKDLSLLQA 178
QY 181 LRESQSYLVEDLERS 195
DB 179 LKEIHPYLVMDLEQS 193

RESULT 10

US-09-841-879B-2
; Sequence 2, Application US/09841879B
; Patent No. US20020142979A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-130001
; CURRENT APPLICATION NUMBER: US/09/841,879B
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: US 09/728,721
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-841-879B-2

Query Match 69.1%; Score 684; DB 10; Length 193;
Best Local Similarity 71.8%; Pred. No. 2.6e-65;
Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;

QY 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDITDKLVSPY 60
DB 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLQMDAIDLTDKLVSY 60
QY 61 LETYGAEALTANVLRDMGLQEMAGOLQAATHOGSGAAPAGIQAPPOSAAKPGHLHFIDQHRA 120
DB 61 LETYGAEALTANVLRDMGLQELAEQLQ-TTKEESGAVAAAAASVPAQSTARTG-HFVDQHRQ 118
QY 121 ALIARVTNVWLLDALYKGKVLTDQYQAVRAEPTNPSPKMRKLFSTPAWNWTKDILLQA 180
DB 119 ALIARVTEVDGVLDAHGSVLTEGQYQAVRAETTSQDKMRKLFSPVPSWNLTKDLSLLQA 178
QY 181 LRESQSYLVEDLERS 195
DB 179 LKEIHPYLVMDLEQS 193

RESULT 11

US-10-295-981-61
; Sequence 61, Application US/10295981
; Publication No. US20030120055A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/10/295,981
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US/09/340,620
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041

```

; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-295-981-61

Query Match
Best Local Similarity 59.1%; Score 684; DB 15; Length 193;
Matches 140; Conservative 19; Mismatches 34; Indels 2; Gaps 2;

QY 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLDTKLVSFY 60
Db 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLQNDALDLDTKLVSY 60

QY 61 LETYGAEALTANVLDMGLQEMAGQLQAATHQSGGAAPAGIQAPPOSAAKPGLHFIDQHRA 120
Db 61 LESYGLTMTVLDMGLQELAEQLQ-TTKESGAVAAAAAASVPAQSTARTG-HFVDQHRQ 118

QY 121 ALIARVTNVWLLDALYKVLTDQEQVAVRAEPTNPSKMRKLFSPPTANNWTKDILLQA 180
Db 119 ALIARVTEVDGVLDAHGSVLTGEGYQAVRAETTSQDKMRKLFSPVPSWNLTCKDSLLQA 178

QY 181 LRESQSYLVEDLERS 195
Db 179 LKEIHPYLVNDLEQS 193

RESULT 12
US-10-106-698-5422
; Sequence 5422, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 5422
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-5422

Query Match
Best Local Similarity 57.0%; Score 564.5; DB 15; Length 158;
Matches 119; Conservative 3; Mismatches 2; Indels 33; Gaps 2;

QY 4 ARDAILDALLENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLDTKLVSFYLET 63
Db 13 ARDAILDALLENLTAEELKKF-----KLVSFYLET 41

QY 64 YGAELTANVLDMGLQEMAGQLQAATHQSGGAAPAGIQAPPOSAAKPGLHFIDQHRAALI 123
Db 42 YGAELTANVLDMGLQEMAGQLQAATHQSGGAAPAGIQAPPOSAAKPGLHFIDQHRAALI 101

QY 124 ARVTNVWLLDALYKVLTDQEQVAVRAEPTNPSKMR 160
Db 102 ARVTNVWLLDALYKVLTDQEQVAVR--PSPPTQAR 136

RESULT 13
US-09-925-299-1421
; Sequence 1421, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1421
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-299-1421

Query Match
Best Local Similarity 51.4%; Score 509; DB 9; Length 136;
Matches 108; Conservative 2; Mismatches 11; Indels 8; Gaps 2;

QY 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLDTKLVSFY 60
Db 11 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLDTKLVSFY 70

QY 61 LETYGAEALTANVLDMGLQEMAGQLQAATHQSGGAAPAG-----IQAPPOSAAKPGLHFID 116
Db 71 LETYGAEALTANVLDMGLQEMAGQLQAATHQSGGAAPLGSRPILLSRQPSOACT----LID 126

QY 117 QHRAALIAR 125
Db 127 QHRAALSRR 135

RESULT 14
US-09-925-299-1421
; Sequence 1421, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1421
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-299-1421

Query Match
Best Local Similarity 51.4%; Score 509; DB 11; Length 136;
Matches 108; Conservative 2; Mismatches 11; Indels 8; Gaps 2;

QY 1 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLDTKLVSFY 60
Db 11 MGRARDAILDALENLTAEELKKFKLLSVPLREGYGRIPRGALLSMDALDLDTKLVSFY 70

QY 61 LETYGAEALTANVLDMGLQEMAGQLQAATHQSGGAAPAG-----IQAPPOSAAKPGLHFID 116
Db 71 LETYGAEALTANVLDMGLQEMAGQLQAATHQSGGAAPLGSRPILLSRQPSOACT----LID 126
```

QY 117 QHRAALIAIAR 125
 Db 127 QHRAALIAIAR 135

RESULT 15

US-09-931-071-7
 ; Sequence 7, Application US/09931071
 ; Patent No. US20020128219A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bertin, John
 ; APPLICANT: Alnemri, Emad S...
 ; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
 ; FILE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
 ; FILE REFERENCE: 07334-335001
 ; CURRENT APPLICATION NUMBER: US/09/931,071
 ; CURRENT FILING DATE: 2002-03-18
 ; PRIOR APPLICATION NUMBER: 09/428,252
 ; PRIOR FILING DATE: 1999-10-27
 ; NUMBER OF SEQ ID NOS: 11
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 7
 ; LENGTH: 90
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-931-071-7

Query Match 47.4%; Score 469; DB 10; Length 90;
 Best Local Similarity 100.0%; Pred. No. 1.2e-42;
 Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 106 SAAKPGHFDQHRAALIAIARVTNVEWLLDALYGVKVLTDQYQAVRAEPTNPSKQKLFSP 165
 Db 1 SAAKPGHFDQHRAALIAIARVTNVEWLLDALYGVKVLTDQYQAVRAEPTNPSKQKLFSP 60
 QY 166 TPANNWTKDQLLQALRESQSYLVEDLERS 195
 Db 61 TPANNWTKDQLLQALRESQSYLVEDLERS 90

Search completed: January 29, 2004, 13:54:38
 Job time : 34.9809 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 29, 2004, 02:57:41 ; Search time 19587.5 Seconds
(without alignments)
11370.118 Million cell updates/sec

Title: US-09-996-617-1
Perfect score: 5444
Sequence: 1 gccccaggcctggagaggt.....aggaataagaagtctac 5444

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pin:*
- 35: em_htg_rod:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	5444	100.0	5444	9	AB023143	AB023143 Homo sapi
2	5400.8	99.2	5404	9	AF310105	AF310105 Homo sapi
3	4882.4	89.7	5100	6	AX459865	AX459865 Sequence
4	4882.4	89.7	5100	9	AF298548	AF298548 Homo sapi
5	4765.6	87.5	5122	6	AX207661	AX207661 Sequence
6	4290	78.8	4290	9	AF229059	AF229059 Homo sapi
7	4281.8	78.7	5306	9	BC051787	BC051787 Homo sapi
8	4148	76.2	4422	6	AX089761	AX089761 Sequence
9	4148	76.2	4422	9	AF229060	AF229060 Homo sapi
10	4100	75.3	4200	9	AF229061	AF229061 Homo sapi
11	4078	74.9	4194	6	AX089763	AX089763 Sequence
12	3958	72.7	4332	9	AF229062	AF229062 Homo sapi
13	3952	72.6	4329	6	AX089765	AX089765 Sequence
14	3807.8	69.9	4123	9	HSM800983	AL117470 Homo sapi
15	3400.4	62.5	4556	6	AX089769	AX089769 Sequence
16	3210.4	59.0	4466	6	AX089771	AX089771 Sequence
17	3094.8	56.8	3419	9	AK026393	AK026393 Homo sapi
18	1710.8	31.4	152812	2	AC090536	AC090536 Homo sapi
19	1710.8	31.4	176730	9	AC055839	AC055839 Homo sapi
20	1566.4	28.8	2373	9	AK026398	AK026398 Homo sapi
21	991.8	18.2	11862	2	HSAC001237	AC001237 Homo sapi
22	954	17.5	11862	2	HSAC001237	AC001237 Homo sapi
23	742.8	13.6	153021	10	AL663042	AL663042 Mouse DNA
24	738	13.6	174564	2	AC121704	AC121704 Rattus no
25	738	13.6	250275	2	AC095695	AC095695 Rattus no
26	728	13.4	221540	2	AC026912	AC026912 Mus muscu
27	722	13.3	239196	2	AC127967	AC127967 Rattus no
28	701.6	12.9	96115	10	AL662908	AL662908 Mouse DNA
29	695.2	12.8	180274	2	AC027185	AC027185 Mus muscu
30	603	11.1	254690	2	AC095179	AC095179 Rattus no
31	399.8	7.3	3827	9	AY095146	AY095146 Homo sapi
32	399.8	7.3	3827	9	AY154467	AY154467 Homo sapi
33	399.8	7.3	69124	2	AC069567	AC069567 Mus muscu
34	398.2	7.3	3466	9	AK095460	AK095460 Homo sapi
35	398.2	7.3	3507	9	BC028069	BC028069 Homo sapi
36	398	7.3	3108	6	AX417214	AX417214 Sequence
37	398	7.3	3300	6	AX459869	AX459869 Sequence
38	398	7.3	3563	9	AY116205	AY116205 Homo sapi
39	398	7.3	3731	9	AY116204	AY116204 Homo sapi
40	397.4	7.3	3395	9	AY116206	AY116206 Homo sapi
41	395.8	7.3	3221	9	AY116207	AY116207 Homo sapi
42	392.2	7.2	108555	2	AC079221	AC079221 Mus muscu
43	376	6.9	3828	9	AF427617	AF427617 Homo sapi
44	374.4	6.9	3237	9	AY092033	AY092033 Homo sapi
45	373	6.9	4170	9	AF410477	AF410477 Homo sapi

ALIGNMENTS

RESULT 1
AB023143
LOCUS
DEFINITION Homo sapiens mRNA for KIAA0926 protein, complete cds.
ACCESSION AB023143
VERSION AB023143.1 GI:4589483
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (sites)
Nagase,T., Ishikawa,K., Suyama,M., Kikuno,R., Hirose,M.,
Miyajima,N., Tanaka,A., Kotani,H., Nomura,N. and Ohara,O.
TITLE Prediction of the coding sequences of unidentified human genes.

5444 bp mRNA linear PRI 16-JUN-1999

XIII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

92246063
10231032
2 (bases 1 to 5444)
Ohtsuka, O., Nagase, T. and Kikuno, R.
Direct Submission
Submitted (04-FEB-1999) Osamu Ohtsuka, Kazusa DNA Research Institute,
Laboratory of DNA Technology, Yana 1532-3, Kisarazu, Chiba
292-0812, Japan (E-mail:cdmainf@kazusa.or.jp, Tel:+81-438-52-3913,
Fax:+81-438-52-3914)
Location/Qualifiers

FEATURES

source

1. .5444
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="hh02962"
/sex="male"
/tissue type="brain"
/clone_lib="pBluescriptII SK plus"
/dev_stage="adult"
1. .5444
/gene="KIAA0926"
523. .4812
/gene="KIAA0926"
/codon_start=1
/product="KIAA0926 protein"
/protein_id="BAA76770.1"
/db_xref="GI:4589484"
/translation="MAGGAWRLALCYLEPLKKEELKEFOLLANKAHSRSSGETPAQ
PKTSGMEVASLVQYQGRADWALHTWEOGLRSLCAQAGAGHSPSPYSPSE
PHLGSPQSTSTAVLMPWHELPACTQSGSERVLRQLDTSGRWRBELSALLYOAL
PSPHESQSPNAPTSTAVLWSPQSPSLAPQEAFTQWPLDSTSGIYTE
IREREREKSEKPNAAVVGTPQARHSLQPHHPWPSRESLCTWPMKNEFNQ
KFTQLLLQRPSPQDPLKRSWPDYVEENHLEIETDLFGPLDQEPRIVLQ
AAGIKGTARQKAWGRLQYDRFVHFVFSRELAQSKVSLAEILKGDGTAP
APIROLSPERLLFLDQVDPVLPQSPSELCLHWSQPODALGLSLGKTLIP
EASFLITARTALONLPSLEQARWVLPQSPSELCLHWSQPODALGLSLGKTLIP
SNKELWALCLVFWNLACTCLMQMKREKLTLSKTTTTLCLHYLAQALQAPLGP
QURDLCSLAAGIWKQKTLFSPDRLKHLGDGAILSTPLKMGILQEHPLISYFIHL
CFQEFPAASVYVLEDEKGRHSNCTIILEKTLEAYGHGLFGASTTFLGLLSDG
EREMENIPIHRLSQNRNMQWVPLQLLIQPHSLHCLYETRNKTLFTQVMAHFE
MCMGVTEMLLVCTFCIFSRHVKKLQLEGRQHSRTWSPTMVLFRWVPTDAYWO
ILPSVLKVTNKLKELDLGNSLSHSAVSKLCTLRPRCLLETLAAGCLTAEDCKD
LAFGRANGLTELDLSFNVLTDAGAKHLQRLQSPCKLQRLVLSGLTSDCCQDL
AVLSASPSLKLDDIQNNLDDVGLRICEGIRHPACKLIRIGLDOTTLSDMRQEL
ALEQKPLLIIFSRKPSVMTPEGLDTGEMSNSTSLKQRLQSGERASHVAQNALK
LLDVSKIPFIABIAEBSSEVPVVELLCVPSASQDLHTKPLGTDDDFWGPVAT
EVVDKKNLYRVHFPVAGSRWPTGLCFVMEAVTEIEFCVMDQFLGEINPOHSM
VAGPLDDIAKEGAVEAVHLPHFVALQGHVDTSLFQMAHFKKEGMLLEKPARVELH
IVLENPSFPLGLVLMHNAIRFIPVTSVLLIYHRHVEEPTFFLYLIPSDCSIRKE
LELCYRSPGEDQLSEFYVGHLSGIRLQVKKDDETLVMEALVPRGDMPLATTIPP
ARIAVPSPLDAPQLLHFVDQYREQLIARTVSEVLDLKLHGVLSQEQYERVLAEATR
PSQMRKLPSLSQSWDRCKDGLYQALKETHPHLIMELWEKSGSKGLLPLSS"

gene
CDS

BASE COUNT
ORIGIN

Query Match 100.0%; Score 5444; DB 9; Length 5444;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCCAGGCGCTGGAGAGTCTTGAAGAAACCTGGAGCCAGAGCCCGGGGCTCCACTCT 60
DB 1 GCGCCAGGCGCTGGAGAGTCTTGAAGAAACCTGGAGCCAGAGCCCGGGGCTCCACTCT 60
QY 61 GGGTCTGAAGGCCATTCCTGCTCTGGGCTCTCCACCCCACTCTTCTCAGCCTT 120
DB 61 GGGTCTGAAGGCCATTCCTGCTCTGGGCTCTCCACCCCACTCTTCTCAGCCTT 120
QY 121 GCAGTCAAGGGTGTATCTCAGAGTCCAGGACCCAGGAGGAGGAATCTGAGGAACA 180
DB 121 GCAGTCAAGGGTGTATCTCAGAGTCCAGGACCCAGGAGGAGGAATCTGAGGAACA 180

181 CAGAACAGTGAAGCGTTGCCACACACCCCATCTCCCGTACCACATCTCCCTCACCTCAC 240
181 CAGAACAGTGAAGCGTTGCCACACACCCCATCTCCCGTACCACATCTCCCTCACCTCAC 240
241 CTTCCCTGCTGGCGCTCGAACCCCATCCAGAGACCTCCCTATCAGTCACTCTTCCAGT 300
241 CTTCCCTGCTGGCGCTCGAACCCCATCCAGAGACCTCCCTATCAGTCACTCTTCCAGT 300
301 GTCTTGACAGGCGCTCTGGGCTCTCCCTCCCTGGCTTTTCCCTACCACTCCCTCTCTAT 360
301 GTCTTGACAGGCGCTCTGGGCTCTCCCTCCCTGGCTTTTCCCTACCACTCCCTCTCTAT 360
361 GGGCGTCTATCTGTAGTGGCTCTGGGATTTATAAACTGGGTTCCGAATGCTGAATAAGA 420
361 GGGCGTCTATCTGTAGTGGCTCTGGGATTTATAAACTGGGTTCCGAATGCTGAATAAGA 420
421 GACGTAAGAGCAAGGCAAGGAGGACAGCACTGTTCTTCCCTGCTGCTGATACCTCACCA 480
421 GACGTAAGAGCAAGGCAAGGAGGACAGCACTGTTCTTCCCTGCTGCTGATACCTCACCA 480
481 CTGGGAAATCATCCCCCAGACACCTCTTAACTCCGGGACAGAGATGGCTGGCGGAGCTGG 540
481 CTGGGAAATCATCCCCCAGACACCTCTTAACTCCGGGACAGAGATGGCTGGCGGAGCTGG 540
541 GGGCGCTGGCGCTGTTACTTGGAGTTCCTGAAGAGGAGGAGCTGAAGGAGTTCAGGCTT 600
541 GGGCGCTGGCGCTGTTACTTGGAGTTCCTGAAGAGGAGGAGCTGAAGGAGTTCAGGCTT 600
601 CTGCTCGCAATAAAGCGCACTCCAGGAGCTCTTGGGTGAGACACCCGCTCAGCCAGAG 660
601 CTGCTCGCAATAAAGCGCACTCCAGGAGCTCTTGGGTGAGACACCCGCTCAGCCAGAG 660
661 AAGACAGTGGCATGGAGTGGCTCTGTAAGTGGTGGCTCAGTATGGGAGGAGCGGGCC 720
661 AAGACAGTGGCATGGAGTGGCTCTGTAAGTGGTGGCTCAGTATGGGAGGAGCGGGCC 720
721 TGGGACCTTAGCCCTCCATACCTGGGAGCAGATGGGCTGAGGTCACTGCTGGCGCCAGCC 780
721 TGGGACCTTAGCCCTCCATACCTGGGAGCAGATGGGCTGAGGTCACTGCTGGCGCCAGCC 780
781 CAGGAAGGGGAGGCGCACTCTCCCTCATTCCTCCATACGCCCAAGTGAACCCCACTGGG 840
781 CAGGAAGGGGAGGCGCACTCTCCCTCATTCCTCCATACGCCCAAGTGAACCCCACTGGG 840
841 TCTCCAGCGAACCCACCTCCACCGCAGTGTCTTCCCTGATCCATGAATTCGCGCGG 900
841 TCTCCAGCGAACCCACCTCCACCGCAGTGTCTTCCCTGATCCATGAATTCGCGCGG 900
901 GGGTGCACCCAGGGCTCAGAGAGAGGCTTTTGGAGACAGCTGCCTGACACATCTGGAAG 960
901 GGGTGCACCCAGGGCTCAGAGAGAGGCTTTTGGAGACAGCTGCCTGACACATCTGGAAG 960
961 CGCTGGAGAGAAATCTCTGCTCTCACTCTCTACCAAGCTCTTCCAAAGCTCTCCCAAGCAT 1020
961 CGCTGGAGAGAAATCTCTGCTCTCACTCTCTACCAAGCTCTTCCAAAGCTCTCCCAAGCAT 1020
1021 GAGTCTCAAGCCAGGAGTCAACCCAGCGCCACATCCACAGAGTGTCTGGGAGCTGG 1080
1021 GAGTCTCAAGCCAGGAGTCAACCCAGCGCCACATCCACAGAGTGTCTGGGAGCTGG 1080
1081 GGATCCCCCAGCTCAGCCCGCTAGCACCCAGAGAGGAGGCTCTTGGAGCCCAATGG 1140
1081 GGATCCCCCAGCTCAGCCCGCTAGCACCCAGAGAGGAGGCTCTTGGAGCCCAATGG 1140
1141 CCTCTGAGTGAAGCGTCAAGGAATTTACTACAGAAATCAGAGAAAGAGAGAGAGAGAA 1200
1141 CCTCTGAGTGAAGCGTCAAGGAATTTACTACAGAAATCAGAGAAAGAGAGAGAGAGAA 1200
1201 TCAGAGAAAGGAGCGCCCGCTTGGGAGCGGTGTAGGAAGCGCCCGCCACAGCGGCACACC 1260
1201 TCAGAGAAAGGAGCGCCCGCTTGGGAGCGGTGTAGGAAGCGCCCGCCACAGCGGCACACC 1260

QY 1261 AGCTACAGCCCAACCAACCCATGGAGCCTTCTGTGAGAGAGAGCCTCTGTGTTCCACA 1320
 Db 1261 AGCTACAGCCCAACCAACCCATGGAGCCTTCTGTGAGAGAGAGCCTCTGTGTTCCACA 1320
 QY 1321 TGGCCCTGAAAAATGAGGATTTTAAACCAAAATTCACACAGCTGCTACTCTCAAAAGA 1380
 Db 1321 TGGCCCTGAAAAATGAGGATTTTAAACCAAAATTCACACAGCTGCTACTCTCAAAAGA 1380
 QY 1381 CCTCAACCCAGAGCCCAAGATCCCTGTGTCAAGAGAGCTGGCCTGATTATGTGGAGAG 1440
 Db 1381 CCTCAACCCAGAGCCCAAGATCCCTGTGTCAAGAGAGCTGGCCTGATTATGTGGAGAG 1440
 QY 1441 AATCGAGGACATTTAATTGAGATCAGAGACTTATTTGGCCCAAGGCTGGATACCCAAAGAA 1500
 Db 1441 AATCGAGGACATTTAATTGAGATCAGAGACTTATTTGGCCCAAGGCTGGATACCCAAAGAA 1500
 QY 1501 CCTCGCATAGTCATCTGAGGGGCTGTGGAAATGGGAAATCAACATGCGCCAGGCGAG 1560
 Db 1501 CCTCGCATAGTCATCTGAGGGGCTGTGGAAATGGGAAATCAACATGCGCCAGGCGAG 1560
 QY 1561 GTGAAGGAAGCCTGGGGGAGAGCCAGCTGTATGGGGAACCGCTTCCAGCATGTCTTCTAC 1620
 Db 1561 GTGAAGGAAGCCTGGGGGAGAGCCAGCTGTATGGGGAACCGCTTCCAGCATGTCTTCTAC 1620
 QY 1621 TTCAAGCTGAGAGAGCTGGCCCAAGTCCAAAGGTGGTGAAGTCTCGCTGAGCTCATCGGAAA 1680
 Db 1621 TTCAAGCTGAGAGAGCTGGCCCAAGTCCAAAGGTGGTGAAGTCTCGCTGAGCTCATCGGAAA 1680
 QY 1681 GATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCCTGTCTAGGCCAGAGCGGTGCTC 1740
 Db 1681 GATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCCTGTCTAGGCCAGAGCGGTGCTC 1740
 QY 1741 TTCAATCCTCGATGGTGTAGATGAGCCAGAGTGGTCTTCAGAGGCCAGAGTCTTAGAGCTC 1800
 Db 1741 TTCAATCCTCGATGGTGTAGATGAGCCAGAGTGGTCTTCAGAGGCCAGAGTCTTAGAGCTC 1800
 QY 1801 TGTCTGCACTGGAGCCAGCCACAGCCGGCGGATGCACTGCTGGGCAAGTTTGTGGGAAA 1860
 Db 1801 TGTCTGCACTGGAGCCAGCCACAGCCGGCGGATGCACTGCTGGGCAAGTTTGTGGGAAA 1860
 QY 1861 ACTATACTTCCGAGGATCCTTCTGTATGATCAAGGCTCGAACACAGCTCTGAGAACCTC 1920
 Db 1861 ACTATACTTCCGAGGATCCTTCTGTATGATCAAGGCTCGAACACAGCTCTGAGAACCTC 1920
 QY 1921 ATTCTCTTTTGGAGCAGCAGCTTGGGTAGAGTCTCGGGTCTCTGAGTCCAGCAGG 1980
 Db 1921 ATTCTCTTTTGGAGCAGCAGCTTGGGTAGAGTCTCGGGTCTCTGAGTCCAGCAGG 1980
 QY 1981 AAGGAAATATTTCTACAGATATTTACAGATGAAGGCAAGCAATTAGAGCCTTTAGTTTG 2040
 Db 1981 AAGGAAATATTTCTACAGATATTTACAGATGAAGGCAAGCAATTAGAGCCTTTAGTTTG 2040
 QY 2041 GTCAAAATCAAAAGAGCTGTGGGCCCTGTGTCTTGTGGCCCTGGGTGTCTGGCTGGCC 2100
 Db 2041 GTCAAAATCAAAAGAGCTGTGGGCCCTGTGTCTTGTGGCCCTGGGTGTCTGGCTGGCC 2100
 QY 2101 TGCACTTGGCTGATGAGCAGATGAAGCGGAGGAAAGCACTACCTGACTTCCAGAGCC 2160
 Db 2101 TGCACTTGGCTGATGAGCAGATGAAGCGGAGGAAAGCACTACCTGACTTCCAGAGCC 2160
 QY 2161 ACCACAAACCTCTGTCTACATTAACCTTGGCCAGGCTCTCAAGCTCAGGCAATTGGAGCC 2220
 Db 2161 ACCACAAACCTCTGTCTACATTAACCTTGGCCAGGCTCTCAAGCTCAGGCAATTGGAGCC 2220
 QY 2221 CAGCTCAGAGACCTCTGCTCTCTGGTGTGAGGGCATCTGGCAAAAAAGACCCCTTTTC 2280
 Db 2221 CAGCTCAGAGACCTCTGCTCTCTGGTGTGAGGGCATCTGGCAAAAAAGACCCCTTTTC 2280
 QY 2281 AGTCCAGATGACCTCAGGAGCATGGGTAGATGGGGCCATCATCTCCACCTCTTTGAAG 2340
 Db 2281 AGTCCAGATGACCTCAGGAGCATGGGTAGATGGGGCCATCATCTCCACCTCTTTGAAG 2340
 QY 2341 ATGGGTATTCTTCAAGAGCACCCCATCCCTCTGAGCTACAGCTTCATTACCTCTGTTTC 2400

Db 2341 ATGGGTATTCTTCAAGAGCACCCCATCCCTCTGAGCTACAGCTTCATTACCTCTGTTTC 2400
 QY 2401 CAAGAGTTCTTTTGAGCAATGTCTCTATGTCTTTGAGGATGAGAGGGGAGAGGTAACAT 2460
 Db 2401 CAAGAGTTCTTTTGAGCAATGTCTCTATGTCTTTGAGGATGAGAGGGGAGAGGTAACAT 2460
 QY 2461 TCTAATTCATCATAGATTTGGAAGAGACGCTAGAGCAATATGGAATATCATGGCCTGTTT 2520
 Db 2461 TCTAATTCATCATAGATTTGGAAGAGACGCTAGAGCAATATGGAATATCATGGCCTGTTT 2520
 QY 2521 GGGGCATCAACACACGTTTCTTATTTGGCCCTGTAAATGATGAGGGGAGAGAGATG 2580
 Db 2521 GGGGCATCAACACACGTTTCTTATTTGGCCCTGTAAATGATGAGGGGAGAGAGATG 2580
 QY 2581 GAGAACATCTTTTCACTGCGGCTGTCTCAGGGGAGAACTTGAATGAGTGGGTCCCGTCC 2640
 Db 2581 GAGAACATCTTTTCACTGCGGCTGTCTCAGGGGAGAACTTGAATGAGTGGGTCCCGTCC 2640
 QY 2641 CTGCAAGTCTGCTGAGCCACACTCTCTGAGTCCCTCCAATGTTGTAGAGACTCGG 2700
 Db 2641 CTGCAAGTCTGCTGAGCCACACTCTCTGAGTCCCTCCAATGTTGTAGAGACTCGG 2700
 QY 2701 AACAAAACTTCTCTGACACAAAGTATGSCCCATTTTGAAGAAATGGGCACTGTGTAGAA 2760
 Db 2701 AACAAAACTTCTCTGACACAAAGTATGSCCCATTTTGAAGAAATGGGCACTGTGTAGAA 2760
 QY 2761 ACAGACATGGAGCTCTTATGTGCACTTTTCTGCAATTAATTCAGCGCCACCTGTAAGAG 2820
 Db 2761 ACAGACATGGAGCTCTTATGTGCACTTTTCTGCAATTAATTCAGCGCCACCTGTAAGAG 2820
 QY 2821 CTTCAAGCTGATTGAGGGCAGGACACAGATCAAATGAGAGCCCAACCATGTAAGTCTG 2880
 Db 2821 CTTCAAGCTGATTGAGGGCAGGACACAGATCAAATGAGAGCCCAACCATGTAAGTCTG 2880
 QY 2881 TTCAAGTGGTCCAGTCAAGATGCTTATGGAGATTTCTTCTCGTCTCAAGGTC 2940
 Db 2881 TTCAAGTGGTCCAGTCAAGATGCTTATGGAGATTTCTTCTCGTCTCAAGGTC 2940
 QY 2941 ACCAGAAACCTGAGAGCTGAGACCTTAAGTGAAACTCGTGAGCACTCTCGAGTGAAG 3000
 Db 2941 ACCAGAAACCTGAGAGCTGAGACCTTAAGTGAAACTCGTGAGCACTCTCGAGTGAAG 3000
 QY 3001 AGTCTTTTGAAGCCTCTGAGAGCCCTCGTCTCTCTGGAGACCTCGGTTGGCTG 3060
 Db 3001 AGTCTTTTGAAGCCTCTGAGAGCCCTCGTCTCTCTGGAGACCTCGGTTGGCTG 3060
 QY 3061 TGTGSCCTCAGCTGAGGACCTGCAAGGACCTTGTGCTTTGGCTGAGAGCCAAACAGACC 3120
 Db 3061 TGTGSCCTCAGCTGAGGACCTGCAAGGACCTTGTGCTTTGGCTGAGAGCCAAACAGACC 3120
 QY 3121 CTGACCGAGCTGGACCTGAGCTTCAATGTGCTCAGGATGCTGGAGCCAAACACCTTTGC 3180
 Db 3121 CTGACCGAGCTGGACCTGAGCTTCAATGTGCTCAGGATGCTGGAGCCAAACACCTTTGC 3180
 QY 3181 CAGAGACTGAGACAGCCAGCTGCAAGCTACAGCACTGCAAGCTGCTGAGCTGTGGCTTC 3240
 Db 3181 CAGAGACTGAGACAGCCAGCTGCAAGCTACAGCACTGCAAGCTGCTGAGCTGTGGCTTC 3240
 QY 3241 ACCTCTGACTGCTGAGGACCTGCGCTCTGTGCTTGTAGTGGCCAGCCCAAGGAG 3300
 Db 3241 ACCTCTGACTGCTGAGGACCTGCGCTCTGTGCTTGTAGTGGCCAGCCCAAGGAG 3300
 QY 3301 CTAGACCTGAGCAGAAACAACTGATGAGTGGCTGGAGTCTCTGTGAGGGGCTC 3360
 Db 3301 CTAGACCTGAGCAGAAACAACTGATGAGTGGCTGGAGTCTCTGTGAGGGGCTC 3360
 QY 3361 AGGCACTCTGCTGCAAACTCATAGCCTGGGGCTGGACAGCAAACTCTGAGTGTATGAG 3420
 Db 3361 AGGCACTCTGCTGCAAACTCATAGCCTGGGGCTGGACAGCAAACTCTGAGTGTATGAG 3420
 QY 3421 ATGAGCAGGAACTGAGGGCCCTGGAGCAGGAGAAACCTCAGCTGCTCATCTTCAGCAGA 3480

Db 3421 ATGAGGACGAACTGAGGCGCTTGGAGCAGGAGAAACCTCAGCTGCTCATCTTTCAGCAGA 3480
 Qy 3481 CGGAACCAAGTGTGATGACCCCTACTGAGGCGCTGGATACGGGAGAGATGATATAGC 3540
 Db 3481 CGGAACCAAGTGTGATGACCCCTACTGAGGCGCTGGATACGGGAGAGATGATATAGC 3540
 Qy 3541 ACATCTCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGCGCTTCCCATCTGTCTCAG 3600
 Db 3541 ACATCTCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGCGCTTCCCATCTGTCTCAG 3600
 Qy 3601 GCTAACTCTCAAACTCTGACGCTGAGCAAGATCTTCCCAATTTCTGAGATTTGCAAGGAA 3660
 Db 3601 GCTAACTCTCAAACTCTGACGCTGAGCAAGATCTTCCCAATTTCTGAGATTTGCAAGGAA 3660
 Qy 3661 AGCTCCCAAGAGTAGTACCGGTGGAACCTTTGTGGGTGCTTCTCTGCTCTCAAGG 3720
 Db 3661 AGCTCCCAAGAGTAGTACCGGTGGAACCTTTGTGGGTGCTTCTCTGCTCTCAAGG 3720
 Qy 3721 GACCTGCATACGAAGCTTTGGGACTGACGATGACTTCTGGGGCCCCACGGGCGCTGTG 3780
 Db 3721 GACCTGCATACGAAGCTTTGGGACTGACGATGACTTCTGGGGCCCCACGGGCGCTGTG 3780
 Qy 3781 GCTACTGAGGTAGTTGACAAAGAAAGAACTTTGTACCGAGTTCACTTCCCTGTAGCTGCG 3840
 Db 3781 GCTACTGAGGTAGTTGACAAAGAAAGAACTTTGTACCGAGTTCACTTCCCTGTAGCTGCG 3840
 Qy 3841 TCCTACCGCTGCCCAACACGGGTCTCTGCTTTGTGATGAGAGAGCGGTGACCGTTGAG 3900
 Db 3841 TCCTACCGCTGCCCAACACGGGTCTCTGCTTTGTGATGAGAGAGCGGTGACCGTTGAG 3900
 Qy 3901 ATTGAATCTGTGTGGGACCACTTCCCTGGGTGAGATCAACCCACAGACAGCTGGATG 3960
 Db 3901 ATTGAATCTGTGTGGGACCACTTCCCTGGGTGAGATCAACCCACAGACAGCTGGATG 3960
 Qy 3961 GTGGAGGSCCTCTGCTGACATCAAGGCTGAGCTGGAGCTGTGGAAGCTGTGACCTC 4020
 Db 3961 GTGGAGGSCCTCTGCTGACATCAAGGCTGAGCTGGAGCTGTGGAAGCTGTGACCTC 4020
 Qy 4021 CCTCATTCTTGGCTTCAAGGGGCCATGTGGACACATCTCTGTTCCAAATGGCCAC 4080
 Db 4021 CCTCATTCTTGGCTTCAAGGGGCCATGTGGACACATCTCTGTTCCAAATGGCCAC 4080
 Qy 4081 TTTTAAAGAGAGGGGATGCTCTCGAAGAGCCAGCCAGGCTGGAGCTGCATCACATAGTT 4140
 Db 4081 TTTTAAAGAGAGGGGATGCTCTCGAAGAGCCAGCCAGGCTGGAGCTGCATCACATAGTT 4140
 Qy 4141 CTGGAAAACCCAGCTTCTCCCTTGGAGTCTCTCGAATGATCATTAATGCCCTG 4200
 Db 4141 CTGGAAAACCCAGCTTCTCCCTTGGAGTCTCTCGAATGATCATTAATGCCCTG 4200
 Qy 4201 CGCTTCATTCCCGTCACTCTGTGTGTGCTTTTACCACCGCTCCATCTGAGGAATC 4260
 Db 4201 CGCTTCATTCCCGTCACTCTGTGTGTGCTTTTACCACCGCTCCATCTGAGGAATC 4260
 Qy 4261 ACCTTCCACTCTACCTGATCCCAAGTACTGCTTCCATTCGGAAGAACTGGAGCTCTGC 4320
 Db 4261 ACCTTCCACTCTACCTGATCCCAAGTACTGCTTCCATTCGGAAGAACTGGAGCTCTGC 4320
 Qy 4321 TATCGAAGCCCTGGAGAACAGCAGCTGTTCTCGGAGTTCTACGTTGGCACTTGGGATCA 4380
 Db 4321 TATCGAAGCCCTGGAGAACAGCAGCTGTTCTCGGAGTTCTACGTTGGCACTTGGGATCA 4380
 Qy 4381 GGGATCAGGCTGCAAGTGAAGCAAGAAAGATGAGACTCTGGTGTGGAGCCCTTGGTG 4440
 Db 4381 GGGATCAGGCTGCAAGTGAAGCAAGAAAGATGAGACTCTGGTGTGGAGCCCTTGGTG 4440
 Qy 4441 AAACACGAGAGATCTCATGCTGCTCAACTACTCTGATCCCTCCAGCCGCGATAGCCGTACCT 4500
 Db 4441 AAACACGAGAGATCTCATGCTGCTCAACTACTCTGATCCCTCCAGCCGCGATAGCCGTACCT 4500
 Qy 4501 TCACCTCTGGATGCCCGCAGTTGTGCACTTTGTGGACCACTGTGAGAGCAGCTGATA 4560
 Db 4501 TCACCTCTGGATGCCCGCAGTTGTGCACTTTGTGGACCACTGTGAGAGCAGCTGATA 4560

Qy 4561 GCCGAGTGACATCGGTGGAGTTGTCTTGGACAAACTGTCATGGACAGGTGCTGAGCCAG 4620
 Db 4561 GCCGAGTGACATCGGTGGAGTTGTCTTGGACAAACTGTCATGGACAGGTGCTGAGCCAG 4620
 Qy 4621 GAGCAGTACGAGAGGGTCTGGCTGAGAAACAGAGGCCACCGCAGATCGGAGAGCTGTT 4680
 Db 4621 GAGCAGTACGAGAGGGTCTGGCTGAGAAACAGAGGCCACCGCAGATCGGAGAGCTGTT 4680
 Qy 4681 AGCTTGACGCTCTGGGACCGGAACTGCAAGATGAGACTCTACCAAGCCCTGAAGGAG 4740
 Db 4681 AGCTTGACGCTCTGGGACCGGAACTGCAAGATGAGACTCTACCAAGCCCTGAAGGAG 4740
 Qy 4741 ACCCATCTCTCACTTATGGAACCTGGGAGAGGCGCAAAAAGGGACTCTCTGCCA 4800
 Db 4741 ACCCATCTCTCACTTATGGAACCTGGGAGAGGCGCAAAAAGGGACTCTCTGCCA 4800
 Qy 4801 CTCAGCAGTGAAGTATCAACACAGCCCTTGAACCTTGAAGTCTGGCTTTGGGTGACCC 4860
 Db 4801 CTCAGCAGTGAAGTATCAACACAGCCCTTGAACCTTGAAGTCTGGCTTTGGGTGACCC 4860
 Qy 4861 TTTCTTGGGTCTCAGTTTCTTCTCTGCAAAAGTTGCGCATCTGGTTCCTTCCAGCA 4920
 Db 4861 TTTCTTGGGTCTCAGTTTCTTCTCTGCAAAAGTTGCGCATCTGGTTCCTTCCAGCA 4920
 Qy 4921 CTAAAGTAATGGAACCTTTGATGATGCTTCTGCGGCAATATGTCCATGCGGAGATG 4980
 Db 4921 CTAAAGTAATGGAACCTTTGATGATGCTTCTGCGGCAATATGTCCATGCGGAGATG 4980
 Qy 4981 CCACAGGGGCCCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT 5040
 Db 4981 CCACAGGGGCCCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT 5040
 Qy 5041 GGCAGACCCCTGCAGACTCATAGAGCTCATCTGTGGGCCACAGCAGCCAGCCTAGA 5100
 Db 5041 GGCAGACCCCTGCAGACTCATAGAGCTCATCTGTGGGCCACAGCAGCCAGCCTAGA 5100
 Qy 5101 GCCCTCGGATCCCATCCAGCGCAAGAGGAAATAGGAGGACATGGAACATTTGCTCCTC 5160
 Db 5101 GCCCTCGGATCCCATCCAGCGCAAGAGGAAATAGGAGGACATGGAACATTTGCTCCTC 5160
 Qy 5161 TGGCTGTCTCAGGGTGGAGCCCAAAATTTGGGTTTACGCTGGAGGCCACGTTGATTC 5220
 Db 5161 TGGCTGTCTCAGGGTGGAGCCCAAAATTTGGGTTTACGCTGGAGGCCACGTTGATTC 5220
 Qy 5221 TGGCTTGTACAGGAAGATCTACAGAGCAAGCCACAGAGTAAAGTGGAGGAGTTT 5280
 Db 5221 TGGCTTGTACAGGAAGATCTACAGAGCAAGCCACAGAGTAAAGTGGAGGAGTTT 5280
 Qy 5281 ATTCAAGAAATAAAGGAGTATCAGAGCTCTTTTGAATTTTCTAGCAGGCTTTCCAGTT 5340
 Db 5281 ATTCAAGAAATAAAGGAGTATCAGAGCTCTTTTGAATTTTCTAGCAGGCTTTCCAGTT 5340
 Qy 5341 TTTTACAGAAACCCCTATAAAATTTTAAATTTTCTTAAATTTTAAAGTTTAAAGTTT 5400
 Db 5341 TTTTACAGAAACCCCTATAAAATTTTAAATTTTCTTAAATTTTAAAGTTTAAAGTTT 5400
 Qy 5401 ACAAAGAAAGAAATGAAATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGAT 5444
 Db 5401 ACAAAGAAAGAAATGAAATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGATTAAGGAT 5444

RESULT 2
 AF310105 LOCUS AF310105 5404 bp mRNA linear PRI 04-MAY-2001
 DEFINITION Homo sapiens NALP1 mRNA, complete cds.
 ACCESSION AF310105
 VERSION AF310105.1 GI:11096302
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 5404).
 Martinon, F., Hofmann, K. and Tschoopp, J.
 The pyrin domain: a possible member of the death domain-fold family
 implicated in apoptosis and inflammation
 Curr. Biol. 11 (4), R118-R120 (2001)
 21148093
 MEDLINE
 PUBMED
 11250163
 2 (bases 1 to 5404)
 Martinon, F., Hofmann, K. and Tschoopp, J.
 Direct Submission
 Submitted (28-SEP-2000) Institute of Biochemistry, University of
 Lausanne, Ch des Boveresses 155, Epalinges 1066, Switzerland
 Location/Qualifiers
 1..5404
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 519..4808
 /note="contains NACHT, LRR and PYD domains"
 /codon_start=1
 /product="NALP1"
 /protein_id="AAG30288.1"
 /db_xref="GI:11096303"
 /translation="MAGGAWRLACVLEFLKKEELKEFOLLANKAHSSSSGTPAQ
 PKTSMESVASYLVAGYQRAWDLALHTWEQMLRSICAQAEAGHSPSPSPSE
 PHLSPSPSTAVLMPWTHLPAAGCTQSGSERVLRQLPDTSGRWRREISASLYQAL
 PSSPHSPSPSTAVLMPWTHLPAAGCTQSGSERVLRQLPDTSGRWRREISASLYQAL
 IREERKESEKRPWAAVVGTPPOAHTSLOPHHPWEPVSRESLCSWPMKNEPNO
 KETQLLLORPHRSPDPLVRSWPDYVEENRHLIEIRDLPGLDTPQEPRIVLQGP
 AAGIKSTUARQKEAWGQYLDYRQFHYFYSRELAQSKVSLAEILGKDGATP
 APIRQLSPERLLFLDGVDFWGLVQPSSELCHWSQFQPADLLGSLGKTLIP
 EASFLITARTLQNLIPSEARWVEVLFQPSSESKKEYFYFYFDEQAIARFLVK
 SNKELWALCLVPWSWLAICTLMOOKRKEKLTLSKTTTLCLHVAQALQAPLGP
 QLRDLCSLAEGWOKKTLFSPDDLKRGDLGAILSTFLKMGILQEHPIPLSYSIHL
 CQPEPAMSYLEDEKGRGKSNCLIDLEKLEAYGHGLFGASTRELLGLLSDG
 EMEENIFRQSQNLNMQWVPSIQLLQPHSLHSLCLYETRNKTLFTQVMAHFE
 MCMVETDMLLYTCIFSRHVKQLQIEGRHRSWTSYTMFLRWVPTDAYWQ
 ILFVLKRVNLTLELDSNLGSLSHGAKSLCKTLRPRCLLETLRAGLTAEDCKD
 LAFLGRANOTLLELDSNLGSLSHGAKSLCKTLRPRCLLETLRAGLTAEDCKD
 ASVLASPSLKLIDLOONLDDVGLRLCEGRHHPACKLIRGLDQTLTSDMRQELR
 ALQEPKQLLISRRKPSVMTPEGLDGTGEMNSSTSLKRLGSRERAAHVAQNLK
 LLDVSKIPIAEIAEESPVEVVELLCVPSPASQDHLTKPLDIDDDFWGTPVAT
 EVDDEKNLYRHFVAGSYRNPENTGLCFWREAVTEIEFCVWPQFLGEINPQSHW
 VAGPLDIDKAEPGVAEVAHPLHFVALQGHVDTSLFQMAHFKEEGMLEKPARVELHH
 IVLENPSPPLGVLLKMIHNLRFIPVTSVLLVHRVPEEVTFLYLIPSDCSIRKE
 LELCVSPGEDQLFSEFYVGHLSGRLQVLDKQDETLVWEALVKEGDLMPATLIPP
 ARIAPSPDAPQLLHFVQYREQLIARTVSEVLDLKLHGVLSQEQYERVAENR
 PSQMRKLFSLSSQSDRKCKDGLYQALKETHPLIMELWEKSKGLLLPSS"
 BASE COUNT 1284 a 1518 c 1467 g 1135 t
 ORIGIN

Query Match 99.2%; Score 5400.8; DB 9; Length 5404;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 5402; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 CAGGGCTGAGAGGCTGAAGAAAACCTGGAGCCAGCAGCCGGGGCTCCACTCTGGGT 64
 DB 1 CAGGGCTGAGAGGCTGAGAAAACCTGGAGCCAGCAGCCGGGGCTCCACTCTGGGT 60

QY 65 TCTGAAGCCCAATTCCTGCTCTGGGGTCTCTCCACCCCACTCTCTTCTCAGCCTTGCAG 124
 DB 61 TCTGAAGCCCAATTCCTGCTCTGGGGTCTCTCCACCCCACTCTCTTCTCAGCCTTGCAG 120

QY 125 CTCAGGGTGTATCTCAGGAGTCCAGACCCAGGAGAGGAGAAATCTCAGGAACACAGA 184
 DB 121 CTCAGGGTGTATCTCAGGAGTCCAGACCCAGGAGAGGAGAAATCTCAGGAACACAGA 180

QY 185 ACAGTGAGGCTTCCACACACCCCATCTCCCGTCACACACATCTCCCGTCACCCCTC 244
 DB 181 ACAGTGAGGCTTCCACACACCCCATCTCCCGTCACACACATCTCCCGTCACCCCTC 240

QY 245 CTTGCTGGCCCTGGAGCCCATCCAGGACCTCCCTATCAGCTGACTTCTTCCAGTGTCT 304

DB 241 CTTGCTGGCCCTGGAGCCCATCCAGGACCTCCCTATCAGCTGACTTCTTCCAATGTCT 300
 QY 305 TGCAGGCCCTCTGGGGTCTCCCTCCCTCCCTCCCTCCCTTCTTCTTCTTCTTCTTCTTCT 364
 DB 301 TGCAGGCCCTCTGGGGTCTCCCTCCCTCCCTCCCTCCCTTCTTCTTCTTCTTCTTCT 360
 QY 365 GTCTATCTGTAGTGGCTCCCTGGGATTTTATAAATCTGGGTTCCGAATGCTGATTAAGAGAG 424
 DB 361 GTCTATCTGTAGTGGCTCCCTGGGATTTTATAAATCTGGGTTCCGAATGCTGATTAAGAGAG 420
 QY 425 GTAAGAGCCCAAGGAG 484
 DB 421 GTAAGAGCCCAAGGAG 480
 QY 485 GAAACATCCCCAGAGACACCCCTCTTAACTCCGGGACAGAGATGGCTGGCGAGAGCTTGGGCC 544
 DB 481 GAAACATCCCCAGAGACACCCCTCTTAACTCCGGGACAGAGATGGCTGGCGAGAGCTTGGGCC 540
 QY 545 GCCTGGCTGTATCTTGGAGTTCCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 604
 DB 541 GCCTGGCTGTATCTTGGAGTTCCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
 QY 605 TCGCCAAATAAAGCGCACTCCAGGAGCTCTTGGGTGAGACACCCCTCAGCCAGAGAGAGA 664
 DB 601 TCGCCAAATAAAGCGCACTCCAGGAGCTCTTGGGTGAGACACCCCTCAGCCAGAGAGAGA 660
 QY 665 CGAGTGGCATGAGGTGGCTCTGTAACCTGGTGGCTCAGTATGGGAGAGAGAGAGAGAGAGAG 724
 DB 661 CGAGTGGCATGAGGTGGCTCTGTAACCTGGTGGCTCAGTATGGGAGAGAGAGAGAGAGAG 720
 QY 725 ACCTAGCCCTCCATACCTGGGAG 784
 DB 721 ACCTAGCCCTCCATACCTGGGAG 780
 QY 785 AAGGGGAGGAGGACCTCTCCCTCATTTCCCTCAGACCCCAAGTGAAACCCACCTTGGGGTCTC 844
 DB 781 AAGGGGAGGAGGACCTCTCCCTCATTTCCCTCAGACCCCAAGTGAAACCCACCTTGGGGTCTC 840
 QY 845 CCAGCAAAACCCACCTCCACCGCAGTGCTAAATGCCCTGGATCCATGAATTCGCGCGGGGT 904
 DB 841 CCAGCAAAACCCACCTCCACCGCAGTGCTAAATGCCCTGGATCCATGAATTCGCGCGGGGT 900
 QY 905 GCACCCAGGGCTCAG 964
 DB 901 GCACCCAGGGCTCAG 960
 QY 965 GGAGAGAAATCTCTGCCTCACTCCTCTACCAAGCTCTTCCAAAGCTCCCCAGACCATGAT 1024
 DB 961 GGAGAGAAATCTCTGCCTCACTCCTCTACCAAGCTCTTCCAAAGCTCCCCAGACCATGAT 1020
 QY 1025 CTCCAAGCCAGGAGTCAACCAAGCCCCACATCCACAGCAGTGTGGGAGCTGGGAT 1084
 DB 1021 CTCCAAGCCAGGAGTCAACCAAGCCCCACATCCACAGCAGTGTGGGAGCTGGGAT 1080
 QY 1085 CCCCACCTCAGCCAGCCCTAGCACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1144
 DB 1081 CCCCACCTCAGCCAGCCCTAGCACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
 QY 1145 TGGATGAAACGTCAGGAATTTACTACAGAAATCAGAGAAATCAGAGAGAGAGAGAGAGAGAGAGAG 1204
 DB 1141 TGGATGAAACGTCAGGAATTTACTACAGAAATCAGAGAAATCAGAGAGAGAGAGAGAGAGAGAGAG 1200
 QY 1205 AGAAAGGAG 1264
 DB 1201 AGAAAGGAG 1260
 QY 1265 TACAGCCCCACCAACCAACCTGGGAGGCTTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1324
 DB 1261 TACAGCCCCACCAACCAACCTGGGAGGCTTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
 QY 1325 CTTGAGAAATGAGGATTTTAAACCAAAATTTACACAGCTGTCTTCTTACAAAGACCTTC 1384
 DB 1321 CTTGAGAAATGAGGATTTTAAACCAAAATTTACACAGCTGTCTTCTTACAAAGACCTTC 1380

1385 ACCCAGAAAGCCAGATCCCTGGTCAAGAGAAAGCTGGCTGATTATGTGGAGGAGAATC 1444
 1381 ACCCAGAAAGCCAGATCCCTGGTCAAGAGAAAGCTGGCTGATTATGTGGAGGAGAATC 1440
 1445 GAGGACATTTAATTGAGATCAGAGACTTATTTGGCCAGGCTGGATACCCAGAAACCTC 1504
 1441 GAGGACATTTAATTGAGATCAGAGACTTATTTGGCCAGGCTGGATACCCAGAAACCTC 1500
 1505 GCATAGTCATCTGCAAGGGGGCTGCTGGAATTTGGGAAGTCAACATGGCCAGGCAAGGTGA 1564
 1501 GCATAGTCATCTGCAAGGGGGCTGCTGGAATTTGGGAAGTCAACATGGCCAGGCAAGGTGA 1560
 1565 AGGAAGCCTGGGGAGAGGCGAGCTGTATGGGAGCCGCTTCCAGCATGCTTCTACTTCA 1624
 1561 AGGAAGCCTGGGGAGAGGCGAGCTGTATGGGAGCCGCTTCCAGCATGCTTCTACTTCA 1620
 1625 GCTGCAGAGAGCTGGCCCAAGTCAAGGTGGTGAAGTCTCGCTGAGCTCATCGGAAAGATG 1684
 1621 GCTGCAGAGAGCTGGCCCAAGTCAAGGTGGTGAAGTCTCGCTGAGCTCATCGGAAAGATG 1680
 1685 GGAAGCCACTCGGCTCCCATTTAGACAGATCTCTGTCTAGGCGAGAGCGGCTGCTTCA 1744
 1681 GGAAGCCACTCGGCTCCCATTTAGACAGATCTCTGTCTAGGCGAGAGCGGCTGCTTCA 1740
 1745 TCCTCGATGGTGTAGATGAGCCAGGATGGTCTTGCAGGAGCGAGTTCCTGAGCTCTGTC 1804
 1741 TCCTCGATGGTGTAGATGAGCCAGGATGGTCTTGCAGGAGCGAGTTCCTGAGCTCTGTC 1800
 1805 TGCACTGGAGCCAGCCAGCCGCGGATGCACTGCTGGGCGAGTTTGGTGGGAAACTA 1864
 1801 TGCACTGGAGCCAGCCAGCCGCGGATGCACTGCTGGGCGAGTTTGGTGGGAAACTA 1860
 1865 TACTTCCGAGGATCCTTCTGATCAGCGCTGGACCAAGCTCTGCAAGACCTCATTC 1924
 1861 TACTTCCGAGGATCCTTCTGATCAGCGCTGGACCAAGCTCTGCAAGACCTCATTC 1920
 1925 CTTCTTTGAGCAGGACGCTGGTGTAGAGTCTGGGGTCTCTGAGTCCAGCAGGAAGG 1984
 1921 CTTCTTTGAGCAGGACGCTGGTGTAGAGTCTGGGGTCTCTGAGTCCAGCAGGAAGG 1980
 1985 AATATTTCTACAGATATTTCAAGATGAAAGCAAGCAATTAAGACCTTTAGTTGGTCA 2044
 1981 AATATTTCTACAGATATTTCAAGATGAAAGCAAGCAATTAAGACCTTTAGTTGGTCA 2040
 2045 AATCAAAACAAAGCTCTGGGCCCTGTCTTGTGGCCCTGGGTGCTGGCTGCA 2104
 2041 AATCAAAACAAAGCTCTGGGCCCTGTCTTGTGGCCCTGGGTGCTGGCTGCA 2100
 2105 CTTGCTGTGATGACAGATGAAAGCAAGCAAACTCACTGACTTCCAGACCACTCA 2164
 2101 CTTGCTGTGATGACAGATGAAAGCAAGCAAACTCACTGACTTCCAGACCACTCA 2160
 2165 CAACCTCTGTCTACATTTACCTTTGCCAGGCTCTCCAAGCTCAGCCATTGGGACCCAGC 2224
 2161 CAACCTCTGTCTACATTTACCTTTGCCAGGCTCTCCAAGCTCAGCCATTGGGACCCAGC 2220
 2225 TCAGAGACCTCTGCTCTGCTGTGCTGAGGCGATCTGGCAAAAAGACCTTTTTCAGTC 2284
 2221 TCAGAGACCTCTGCTCTGCTGTGCTGAGGCGATCTGGCAAAAAGACCTTTTTCAGTC 2280
 2285 CAGATCACCTCAGGAAGCATGGGTAGATGGGGCATCATCTCCACTTCTTGAAGATGG 2344
 2281 CAGATCACCTCAGGAAGCATGGGTAGATGGGGCATCATCTCCACTTCTTGAAGATGG 2340
 2345 GTATTTCTCAAGAGCACCCCATCCCTCTGAGCTTACAGCTTCACTCTGTTTTCCAAAG 2404
 2341 GTATTTCTCAAGAGCACCCCATCCCTCTGAGCTTACAGCTTCACTCTGTTTTCCAAAG 2400
 2405 AGTTCTTTGAGCAATGCTCTATGCTTGGAGGATGAGAGGGGAGGATTAACATTTCTA 2464
 2401 AGTTCTTTGAGCAATGCTCTATGCTTGGAGGATGAGAGGGGAGGATTAACATTTCTA 2460

2465 ATTGCATCATAGATTTTGAAAAGACGCTAGAAAGCATATGGAAATACATGGCCTGTTTGGGG 2524
 2461 ATTGCATCATAGATTTTGAAAAGACGCTAGAAAGCATATGGAAATACATGGCCTGTTTGGGG 2520
 2525 CATCAACACACAGCTTCTATTTGGCCCTGTTAAGTGTAGGGGAGAGAGATGGAGA 2584
 2521 CATCAACACACAGCTTCTATTTGGCCCTGTTAAGTGTAGGGGAGAGAGATGGAGA 2580
 2585 ACATCTTTCACTGCCGGCTGTCTCAGGGGAGGAACCTGATGCAGTGGGTCCCGTCCCTGTC 2644
 2581 ACATCTTTCACTGCCGGCTGTCTCAGGGGAGGAACCTGATGCAGTGGGTCCCGTCCCTGTC 2640
 2645 AGCTGCTGCTGCAGCACACACTCTCTGAGTCCCTCACTGCTGTGTAAGAGACTCGGAACA 2704
 2641 AGCTGCTGCTGCAGCACACACTCTCTGAGTCCCTCACTGCTGTGTAAGAGACTCGGAACA 2700
 2705 AAAAGTCTCTGACACAGTGTATGGCCCATTTTGAAGAAATGGGCATGTGTAGAAAACAG 2764
 2701 AAAAGTCTCTGACACAGTGTATGGCCCATTTTGAAGAAATGGGCATGTGTAGAAAACAG 2760
 2765 ACATGGAGCTCTTATGTGTGCACTTTCTGCATTTAAATTCAGCGGCCACCTGGAAGAGCTTC 2824
 2761 ACATGGAGCTCTTATGTGTGCACTTTCTGCATTTAAATTCAGCGGCCACCTGGAAGAGCTTC 2820
 2825 AGCTGATTTGAGGGCAGGACACAGATCAACATGGAGCCCAACCATGTGTCTGTTTCA 2884
 2821 AGCTGATTTGAGGGCAGGACACAGATCAACATGGAGCCCAACCATGTGTCTGTTTCA 2880
 2885 GGTGGTCCAGTCCAGATGCTTATGGAGATCTTCTTCGCTCTCAAGTCAACA 2944
 2881 GGTGGTCCAGTCCAGTCCAGATGCTTATGGAGATCTTCTTCGCTCTCAAGTCAACA 2940
 2945 GAAACCTCAAGAGCTGGAACCTTAAGTGAACCTCGCTGAGCACTCTGAGTGAAGAGTC 3004
 2941 GAAACCTCAAGAGCTGGAACCTTAAGTGAACCTCGCTGAGCACTCTGAGTGAAGAGTC 3000
 3005 TTTGTAAGACCTCTGAGACGCTCTGCTGCTCTGAGAGACCTCGGTTGGCTGGCTGTG 3064
 3001 TTTGTAAGACCTCTGAGACGCTCTGCTGCTCTGAGAGACCTCGGTTGGCTGGCTGTG 3060
 3065 GCCTCAGAGCTGAGACCTGCAAGGACCTTGGCTTGGCTGAGAGCCCAACAGACCTTGA 3124
 3061 GCCTCAGAGCTGAGACCTGCAAGGACCTTGGCTTGGCTGAGAGCCCAACAGACCTTGA 3120
 3125 CCGAGCTGCACTGAGCTTCAATGTGCTCAGGATGCTGGAGCCAAACCTTTGCCAGA 3184
 3121 CCGAGCTGCACTGAGCTTCAATGTGCTCAGGATGCTGGAGCCAAACCTTTGCCAGA 3180
 3185 GACTGAGACAGCCGAGCTGCAAGCTTACAGCACTGCACTGCTGCTGCTGCTGCTGCTGCT 3244
 3181 GACTGAGACAGCCGAGCTGCAAGCTTACAGCACTGCACTGCTGCTGCTGCTGCTGCTGCT 3240
 3245 CTGACTGCTGCTGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3304
 3241 CTGACTGCTGCTGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300
 3305 ACCTGACGACGAAACCTGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3364
 3301 ACCTGACGACGAAACCTGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3360
 3365 ATCTGCTGCTGCAAACTCATACGCTTGGGCTGGACCACTCTGATGATGATGATGATGATGAT 3424
 3361 ATCTGCTGCTGCAAACTCATACGCTTGGGCTGGACCACTCTGATGATGATGATGATGATGAT 3420
 3425 GGCAGGAACCTGAGGCGCTTGGAGGAGGAAACCTCAGCTGCTCATCTTTCAGCAGCCGA 3484
 3421 GGCAGGAACCTGAGGCGCTTGGAGGAGGAAACCTCAGCTGCTCATCTTTCAGCAGCCGA 3480
 3485 AACCAAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3544
 3481 AACCAAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3540
 3545 CCTCACTCAAGCGGACAGAGACTCGGATCAGAGAGGGGGCTTCCCATGTTGCTCAGGCTA 3604

Db 3541 ||||| CCTCACTAAGCGGACAGACTCGGATCAGAGAGGCGCGCTTCCCATGTGCTCAGGCTA 3600
 QY 3605 ATCTCAAACTCTGTGACGCTGAGCAAGATCTTCCCAATTCCTGAGATGTCAGAGGAAAGCT 3664
 Db 3601 ATCTCAAACTCTGTGACGCTGAGCAAGATCTTCCCAATTCCTGAGATGTCAGAGGAAAGCT 3660
 QY 3665 CCCAGAGGTAGTACCGGTGGAACCTTCTGCGTGCCTTCTCCTCCTCTCAAGGGGACC 3724
 Db 3661 CCCAGAGGTAGTACCGGTGGAACCTTCTGCGTGCCTTCTCCTCCTCTCAAGGGGACC 3720
 QY 3725 TGCATACGAAGCCTTTGGGAGCTGACGATGACTTCTGGGGCCCGACGGGGCTGTGGCTA 3784
 Db 3721 TGCATACGAAGCCTTTGGGAGCTGACGATGACTTCTGGGGCCCGACGGGGCTGTGGCTA 3780
 QY 3785 CTGAGGTAGTGTGACAAAGAAAGAACTTCTGACGAGTTTCACTTCCCTGTAGCTGGCTCT 3844
 Db 3781 CTGAGGTAGTGTGACAAAGAAAGAACTTGTGACGAGTTTCACTTCCCTGTAGCTGGCTCT 3840
 QY 3845 ACCGCTGGCCCAACACGGGTCTCTGCTTTGTGATGAGAAGCGGTGACCGTTGAGATTG 3904
 Db 3841 ACCGCTGGCCCAACACGGGTCTCTGCTTTGTGATGAGAAGCGGTGACCGTTGAGATTG 3900
 QY 3905 AATTCTGTGTGGGACCGATTCCTGGGTGAGATCAACCCACAGCACAGCTGGATGGTGG 3964
 Db 3901 AATTCTGTGTGGGACCGATTCCTGGGTGAGATCAACCCACAGCACAGCTGGATGGTGG 3960
 QY 3965 CAGGGCCTCTGTGGACATCAAGGCTGAGCTGGAGCTGTGGAAGCTGTGCACCTCCCTC 4024
 Db 3961 CAGGGCCTCTGTGGACATCAAGGCTGAGCTGGAGCTGTGGAAGCTGTGCACCTCCCTC 4020
 QY 4025 ACTTTGTGGCTCTCAAGGGGGCCATGTGGACACATCCCTGTTCCAAATGGGCCACTTTA 4084
 Db 4021 ACTTTGTGGCTCTCAAGGGGGCCATGTGGACACATCCCTGTTCCAAATGGGCCACTTTA 4080
 QY 4085 AAGAGAGGGGATGCTCTCGGAGAACCGACGAGCTGGAGCTGTGCATCATAGTTCTGG 4144
 Db 4081 AAGAGAGGGGATGCTCTCGGAGAACCGACGAGCTGGAGCTGTGCATCATAGTTCTGG 4140
 QY 4145 ARAACCCAGCTTCTCCCTTGGGAGTCTCTCGTGAATGATCCATATGCTCCGCT 4204
 Db 4141 ARAACCCAGCTTCTCCCTTGGGAGTCTCTCGTGAATGATCCATATGCTCCGCT 4200
 QY 4205 TCATTCCCGTCACTCTGTGGTGTGCTTTTACACCGCGTCCATCTCGAGGAAGTCACT 4264
 Db 4201 TCATTCCCGTCACTCTGTGGTGTGCTTTTACACCGCGTCCATCTCGAGGAAGTCACT 4260
 QY 4265 TCCACTCTACTGATCCCAAGTGA CTGCTCCATTCGGAAGGAACCTGGAGCTCTGCTATC 4324
 Db 4261 TCCACTCTACTGATCCCAAGTGA CTGCTCCATTCGGAAGGAACCTGGAGCTCTGCTATC 4320
 QY 4325 GAAGCCCTGGAGAACCGAGCTTCTCGGAGTTCTACGTTGGCCACTTGGGATCAGGGA 4384
 Db 4321 GAAGCCCTGGAGAACCGAGCTTCTCGGAGTTCTACGTTGGCCACTTGGGATCAGGGA 4380
 QY 4385 TCAGGCTGCAAGTGAAGAACAGATGAGACTCTGTGTGGGAGGCTTGGTGAAC 4444
 Db 4381 TCAGGCTGCAAGTGAAGAACAGATGAGACTCTGTGTGGGAGGCTTGGTGAAC 4440
 QY 4445 CAGGAGATCTCATGCTTGAACCTGATCTGATCCCTCAGCCCGCATAGCGGTACCTTAC 4504
 Db 4441 CAGGAGATCTCATGCTTGAACCTGATCTGATCCCTCAGCCCGCATAGCGGTACCTTAC 4500
 QY 4505 CTCTGGATGCCCGGAGTCTCTGCACCTTGTGACCAAGTATCAGAGCAGCTGATAGCCC 4564
 Db 4501 CTCTGGATGCCCGGAGTCTCTGCACCTTGTGACCAAGTATCAGAGCAGCTGATAGCCC 4560
 QY 4565 GAGTGACATCGGTGGAGTGTCTTGGACAACTGATGACAGAGTGTGAGCAGGAGC 4624
 Db 4561 GAGTGACATCGGTGGAGTGTCTTGGACAACTGATGACAGAGTGTGAGCAGGAGC 4620
 QY 4625 AGTACGAGAGGTGCTGGCTGAGAACACAGGAGCCCGACGATGCGGAAGCTTTCAGCT 4684

Db 4621 AGTAGAGAGGGTGTGGCTGAGAAACAGAGGCCAGCCAGCATGGGAGAGCTGTTCACT 4680
 QY 4685 TGAGCCAGTCTCTGGACCGGAAGTGCAAAGATGGA CTCTACCAAGCCCTGAAGGAGACCC 4744
 Db 4681 TGAGCCAGTCTCTGGACCGGAAGTGCAAAGATGGA CTCTACCAAGCCCTGAAGGAGACCC 4740
 QY 4745 ATCTCACCTCATTTATGGAACCTCTGGGAGAAAGGAGCAAAAAGGACTCTCTGCCACTCA 4804
 Db 4741 ATCTCACCTCATTTATGGAACCTCTGGGAGAAAGGAGCAAAAAGGACTCTCTGCCACTCA 4800
 QY 4805 GCAGCTGAAGTATCAACACAGCCAGCCCTTCAACCTTGAAGCTCTGGCTTGGCTGACCTTCT 4864
 Db 4801 GCAGCTGAAGTATCAACACAGCCAGCCCTTCAACCTTGAAGCTCTGGCTTGGCTGACCTTCT 4860
 QY 4865 TTGGTCTCAGTCTCTCTCTGCAAAACAAGTTGCCATCTGTTTGGCTTTCAGACATAA 4924
 Db 4861 TTGGTCTCAGTCTCTCTCTGCAAAACAAGTTGCCATCTGTTTGGCTTTCAGACATAA 4920
 QY 4925 AGTAATGGAACCTTTGATGATGCTTGTCTGGGCATTTATGTCATGTCAGGAGTGCAC 4984
 Db 4921 AGTAATGGAACCTTTGATGATGCTTGTCTGGGCATTTATGTCATGTCAGGAGTGCAC 4980
 QY 4985 AGGGGGCCCCCAGTCCAGGTGGCTAAACAGCATCTCAGGGAATGTCCATCTGGAGCTGGCA 5044
 Db 4981 AGGGGGCCCCCAGTCCAGGTGGCTAAACAGCATCTCAGGGAATGTCCATCTGGAGCTGGCA 5040
 QY 5045 AGACCCCTGCAGACCTCATAGAGCCCTCATCTGGTGGCCACACAGCCCAAGCCCTAGAGCCC 5104
 Db 5041 AGACCCCTGCAGACCTCATAGAGCCCTCATCTGGTGGCCACACAGCCCAAGCCCTAGAGCCC 5100
 QY 5105 TCCGATCCCATCCAGGCGCAAGAGGAATAGGAGGACATGGAACCATTTGCCCTCTGGC 5164
 Db 5101 TCCGATCCCATCCAGGCGCAAGAGGAATAGGAGGACATGGAACCATTTGCCCTCTGGC 5160
 QY 5165 TCTGTCCAGGCTGAGCCCCAAAATTGGGTTTTCAGGTTGGAGGACGCTGGATTTCTGG 5224
 Db 5161 TGTGTCCAGGCTGAGCCCCAAAATTGGGTTTTCAGGTTGGAGGACGCTGGATTTCTGG 5220
 QY 5225 CTTTGTACAGGAAGATCTACAGAGCAAGCCAAACAGAGTAAAGTGAAGGAAGTTTATTC 5284
 Db 5221 CTTTGTACAGGAAGATCTACAGAGCAAGCCAAACAGAGTAAAGTGAAGGAAGTTTATTC 5280
 QY 5285 AGAAATTAAGAGGATCACAGCTCTTTTAGAATTTGTCTAGCAGGCTTTCAGTTTATA 5344
 Db 5281 AGAAATTAAGAGGATCACAGCTCTTTTAGAATTTGTCTAGCAGGCTTTCAGTTTATA 5340
 QY 5345 CCAGAAAACCCCTATAAATTAATAATTTTACTTTAAATTTAAGATTAATAAATAACAA 5404
 Db 5341 CCAGAAAACCCCTATAAATTAATAATTTTACTTTAAATTTAAGATTAATAAATAAATAA 5400
 QY 5405 AAAAA 5408
 Db 5401 AAAAA 5404

RESULT 3
 AX459865
 LOCUS 5100 bp DNA linear PAT 08-JUL-2002
 DEFINITION Sequence 8 from Patent WO0240668.
 ACCESSION AX459865
 VERSION AX459865.1 GI:21725641
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 1
 Tschopp, J. and Martinon, F.
 Proteins and dna sequences underlying these proteins used for
 treating inflammations
 Patent: WO 0240668-A 8 23-MAY-2002;
 Apotech Research and Development Ltd. (CH)
 Location/Qualifiers

Db 2133 AGCATATGAAATACATGGGCTGTGTTGGGGCATCAACCAACGTTTCCTATTTGGGCTGT 2192
QY 2556 AAGTCATGAGGGGAGAGAGATGAGAAACATCTTTTCACTGCCGCTCTCTCAGGGAG 2615
Db 2193 AAGTCATGAGGGGAGAGAGATGAGAAACATCTTTTCACTGCCGCTCTCTCAGGGAG 2252
QY 2616 GAACTTGATGAGTGGGTCCGTCCTGAGCTGTGTCGACGACACATCTCTGAGATC 2675
Db 2253 GAACTTGATGAGTGGGTCCGTCCTGAGCTGTGTCGACGACACATCTCTGAGATC 2312
QY 2676 CTTCCATCTGTTGATGAGACTCGGAAACAAAGCTTCCTGACACAGTGCATGGCCATTT 2735
Db 2313 CTTCCATCTGTTGATGAGACTCGGAAACAAAGCTTCCTGACACAGTGCATGGCCATTT 2372
QY 2736 CGAAGAAATGGGCATCTGTGTGAAACACAGACATGGAGCTCTTAGTGTGACATTTCTGCAT 2795
Db 2373 CGAAGAAATGGGCATCTGTGTGAAACACAGACATGGAGCTCTTAGTGTGACATTTCTGCAT 2432
QY 2796 TAAATTCAGCCGCAAGTGAAGCTTCAAGTGAATGAGGGCAGGACACAGATCAAC 2855
Db 2433 TAAATTCAGCCGCAAGTGAAGCTTCAAGTGAATGAGGGCAGGACACAGATCAAC 2492
QY 2856 ATGGAGCCCAACATGGTAGTCTGTTCAAGTGGGTCCAGTCACAGATGCCCTATTGGCA 2915
Db 2493 ATGGAGCCCAACATGGTAGTCTGTTCAAGTGGGTCCAGTCACAGATGCCCTATTGGCA 2552
QY 2916 GATTCCTTTCTCGTCTCAAGTCAACAGAACCTGAAGGAGCTGGACCTTAAGTGAAC 2975
Db 2553 GATTCCTTTCTCGTCTCAAGTCAACAGAACCTGAAGGAGCTGGACCTTAAGTGAAC 2612
QY 2976 CTGCTGAGCCATCTGCGAGTGAAGAGTCTTTGTAAAGCCCTGAGACGCCCTCGCTGCT 3035
Db 2613 CTGCTGAGCCATCTGCGAGTGAAGAGTCTTTGTAAAGCCCTGAGACGCCCTCGCTGCT 2672
QY 3036 CTGAGACCCCTCGGCTTGGTGGCTGTGGCTCAACAGTGAAGCTGCAAGGACCTTGC 3095
Db 2673 CTGAGACCCCTCGGCTTGGTGGCTGTGGCTCAACAGTGAAGCTGCAAGGACCTTGC 2732
QY 3096 CTTTGGGCTGAGAGCCAAACAGACCTGACCGAGCTGGACCTTAAGTGTCTAC 3155
Db 2733 CTTTGGGCTGAGAGCCAAACAGACCTGACCGAGCTGGACCTTAAGTGTCTAC 2792
QY 3156 GGATGCTGGAGCCAAACACCTTTGCGAGACTGAGACGCGGCTGCAAGCTACAGCG 3215
Db 2793 GGATGCTGGAGCCAAACACCTTTGCGAGACTGAGACGCGGCTGCAAGCTACAGCG 2852
QY 3216 ACTGACGTGGTCACTGTGGCTCAAGTGTGACTGCTGCTGAGGACCTGGCTCTGTGCT 3275
Db 2853 ACTGACGTGGTCACTGTGGCTCAAGTGTGACTGCTGCTGAGGACCTGGCTCTGTGCT 2912
QY 3276 TAGTGCCAGCCAGCTGAAGAGCTAGACCTGACGAGAAACAACTGGATGACGCTTGC 3335
Db 2913 TAGTGCCAGCCAGCTGAAGAGCTAGACCTGACGAGAAACAACTGGATGACGCTTGC 2972
QY 3336 CGTGCGACTCTCTGTGAGGGGCTCAGGCATCTGCTGCTGCAAACTCATAGCTGCGGCT 3395
Db 2973 CGTGCGACTCTCTGTGAGGGGCTCAGGCATCTGCTGCTGCAAACTCATAGCTGCGGCT 3032
QY 3396 GGACACAGACAACTCTGAGTGAATGAGTGAAGGACGAACTGAGGGCTCTGGAGCAGAGAA 3455
Db 3033 GGACACAGACAACTCTGAGTGAATGAGTGAAGGACGAACTGAGGGCTCTGGAGCAGAGAA 3092
QY 3456 ACCTGAGCTCTCATCTTACGACGAGCGGAAACCAAGTGTGATGACCCCTACTGAGGCT 3515
Db 3093 GCCTGAGCTCTCATCTTACGACGAGCGGAAACCAAGTGTGATGACCCCTATTGAGGCT 3152
QY 3516 GGATACGGGAGAGATGAGTGAATGAGTGAAGGACGAACTCGGATCAGA 3575
Db 3153 GGATACGGGAGAGATGAGTGAATGAGTGAAGGACGAACTCGGATCAGA 3212
QY 3576 GAGGGCGGCTTCCATGTTGCTCAGGCTAATCTCAAACTCCTGGAAGTGAAGATCTTT 3635
Db 3213 GAGGGCGGCTTCCATGTTGCTCAGGCTAATCTCAAACTCCTGGAAGTGAAGATCTTT 3272

QY 3636 CCCAATTGCTGAGATTGCAGAGAAAGCTCCCGAGAGTAGTACCGGTGGAACCTTTGTG 3695
Db 3273 CCCAATTGCTGAGATTGCAGAGAAAGCTCCCGAGAGTAGTACCGGTGGAACCTTTGTG 3332
QY 3696 CGTGCTTTCTCCTGCTCTCAAGGGGACCTGCATACAGACCTTTGGGACCTGACGATGA 3755
Db 3333 CGTGCTTTCTCCTGCTCTCAAGGGGACCTGCATACAGACCTTTGGGACCTGACGATGA 3392
QY 3756 CTTCTGGGGCCCAACGGGGCTGTGGCTACTGAGGTATTTGACAAAGAAAGAACTTTGA 3815
Db 3393 CTTCTGGGGCCCAACGGGGCTGTGGCTACTGAGGTATTTGACAAAGAAAGAACTTTGA 3452
QY 3816 CCGAGTTCACTTTCCCTGTAGTGGCTCTACCGCTGGGCCCAACACCGGCTCTCTGCTTGT 3875
Db 3453 CCGAGTTCACTTTCCCTGTAGTGGCTCTACCGCTGGGCCCAACACCGGCTCTCTGCTTGT 3512
QY 3876 GATGAGAGAACGGGTGACCGTTGAGATTTGAATTTCTGTGTGGGACCAAGTTCTTGGTGA 3935
Db 3513 GGTGAGAGAACGGGTGACCGTTGAGATTTGAATTTCTGTGTGGGACCAAGTTCTTGGTGA 3572
QY 3936 GATCAACCCACACAGACAGCTGGATGGTGGGAGGCTCTGCTGGACATCAAGGCTGAGCC 3995
Db 3573 GATCAACCCACACAGACAGCTGGATGGTGGGAGGCTCTGCTGGACATCAAGGCTGAGCC 3632
QY 3996 TGGAGCTGTGGAAGCTGTGCACCTCCCTCACTTTTGTGGCTCTCCAAGGGGGCCATGTGGA 4055
Db 3633 TGGAGCCGTGGAGAGCTGTGCACCTCCCTCACTTTTGTGGCTCTCCAAGGGGGCCATGTGGA 3692
QY 4056 CACATCCCTGTTTCCAAATGGGCCCACTTTAAAGAGAGGGGATGCTCTGGAGAACCCAGC 4115
Db 3693 CACATCCCTGTTTCCAAATGGGCCCACTTTAAAGAGAGGGGATGCTCTGGAGAACCCAGC 3752
QY 4116 CAGGCTGAGCTGCATACATAGTTCTGGAACCCAGCTTCTCCCTCTGGAGTCTCT 4175
Db 3753 CAGGCTGAGCTGCATACATAGTTCTGGAACCCAGCTTCTCCCTCTGGAGTCTCT 3812
QY 4176 CTTGAAAATGATCCATAATGCCCTGCGCTTCACTCCCTGCTGTGGTGTGCTTTA 4235
Db 3813 CTTGAAAATGATCCATAATGCCCTGCGCTTCACTCCCTGCTGTGGTGTGCTTTA 3872
QY 4236 CCACCGCTCACTCTGAGGAAAGTCACTTCACTCACTCACTGATCCCAAGTGACTGCTC 4295
Db 3873 CCACCGCTCACTCTGAGGAAAGTCACTTCACTCACTCACTGATCCCAAGTGACTGCTC 3932
QY 4296 CATTCCGAGGAACTGGAGCTCTGCTATCGAAGCCCTGAGAGAACCCAGCTTCTTCGGA 4355
Db 3933 CATTCCGAGGAACTGGAGCTCTGCTATCGAAGCCCTGAGAGAACCCAGCTTCTTCGGA 3992
QY 4356 GTTCTACGTTGGCCCACTTGGGATCAGGGATCAGGCTGCAAGTGAAGAAAGATGA 4415
Db 3993 GTTCTACGTTGGCCCACTTGGGATCAGGGATCAGGCTGCAAGTGAAGAAAGATGA 4052
QY 4416 GACTCTGTGTGGGAGGCTTGGTGAACACAGGAGATCTCATGCTGTCAACTCTCTGAT 4475
Db 4053 GACTCTGTGTGGGAGGCTTGGTGAACACAGGAGATCTCATGCTGTCAACTCTCTGAT 4112
QY 4476 CCCTCCAGCCCATAGCCGTACCTTCACTCTGATGCCCGCAGTTCCTGCACTTGT 4535
Db 4113 CCCTCCAGCCCATAGCCGTACCTTCACTCTGATGCCCGCAGTTCCTGCACTTGT 4172
QY 4536 GGACCAAGTATCGAGACAGCTCATAGCCCGAGTGAATCGGTGGAGGTTGTCTTGGACAA 4595
Db 4173 GGACCAAGTATCGAGACAGCTCATAGCCCGAGTGAATCGGTGGAGGTTGTCTTGGACAA 4232
QY 4596 ACTGCATGGAACAGGTGCTGAGCCAGGACGATGACGAGGGTGTGCTGGTGAAGAACGAG 4655
Db 4233 ACTGCATGGAACAGGTGCTGAGCCAGGACGATGACGAGGGTGTGCTGGTGAAGAACGAG 4292
QY 4656 GCCCAGCCAGATGCGGAAGCTGTTCACTTGAAGCCAGTCTCTGGACCGGAGTGCAGAA 4715
Db 4293 GCCCAGCCAGATGCGGAAGCTGTTCACTTGAAGCCAGTCTCTGGACCGGAGTGCAGAA 4352

QY	4716	TGGACTCTACCAAGCCCTGAAGAGACCCATCTCCTACCTCATTTATGAACTCTGGGAGAA	4775
Db	4353	TGGACTCTACCAAGCCCTGAAGAGACCCATCTCCTACCTCATTTATGAACTCTGGGAGAA	4412
QY	4776	GGCAGCAAAAAGGAGCTCTGCGCATCTGACAGCTGAAGTATCAACACAGGCCCTTGACC	4835
Db	4413	GGCAGCAAAAAGGAGCTCTGCGCATCTGACAGCTGAAGTATCAACACAGGCCCTTGACC	4472
QY	4836	CTTGAGTCTGCGCTTTGGCTGAGCCCTCTTTGGGTCTCGATTTCTTTCTCGCAACAG	4895
Db	4473	CTTGAGTCTGCGCTTTGGCTGAGCCCTCTTTGGGTCTCGATTTCTTTCTCGCAACAG	4532
QY	4896	TTGCCATCTGGTTTGGCTTCCAGCATCTAAAGTAAATGAACTTTGATGATGCTTGGCTGG	4955
Db	4533	TTGCCATCTGGTTTGGCTTCCAGCATCTAAAGTAAATGAACTTTGATGATGCTTGGCTGG	4592
QY	4956	GCATTATGTGTCATGCCAGGAGTGCACAGGGGGCCCGAGTCCAGGTGGCTTAACAGCA	5015
Db	4593	GCATTATGTGTCATGCCAGGAGTGCACAGGGGGCCCGAGTCCAGGTGGCTTAACAGCA	4652
QY	5016	TCTCAGGGAATGTCATCTGGAGCTGGCAAGACCCCTGCAGACCTCATAGACCTCATCT	5075
Db	4653	TCTCAGGGAATGTCATCTGGAGCTGGCAAGACCCCTGCAGACCTCATAGACCTCATCT	4712
QY	5076	GGTGGCCACAGCAGCAAGCCTAGAGCCCTCCGGATCCCATCAGGCGCAAGAGGAATA	5135
Db	4713	GGTGGCCACAGCAGCAAGCCTAGAGCCCTCCGGATCCCATCAGGCGCAAGAGGAATA	4772
QY	5136	GGAGGACATGGAACCATTTGGCTCTGGCTGTGTCTCAGGGTGAAGCCCAAAATTTGGGT	5195
Db	4773	GGAGGACATGGAACCATTTGGCTCTGGCTGTGTCTCAGGGTGAAGCCCAAAATTTGGGT	4832
QY	5196	TCAGCCTGGGAGCCAGCTGGATCTTGGCTTTGTACAGGAATCTACAGAGCAAGCC	5255
Db	4833	TCAGCCTGGGAGCCAGCTGGATCTTGGCTTTGTACAGGAATCTACAGAGCAAGCC	4892
QY	5256	AACAGAGTAAAGTGAAGGAAGTTTATTCAGAAAAATAAGAGGATTCACAGCTCTTTTAG	5315
Db	4893	AACAGAGTAAAGTGAAGGAAGTTTATTCAGAAAAATAAGAGGATTCACAGCTCTTTTAG	4952
QY	5316	AATTCTCTAGCAGGCTTTCCAGTTTTTACCAGAAAAACCCTATATAATTTAAAAATTTT	5375
Db	4953	AATTCTCTAGCAGGCTTTCCAGTTTTTACCAGAAAAACCCTATATAATTTAAAAATTTT	5012
QY	5376	ACTTAATTTAAGAAATTAATAAATAACAAAAAGAAAAATGAATAAAGAAATAGAA	5435
Db	5013	ACTTAATTTAAGAAATTAATAAATAACAAAAAGAAAAATGAATAAAGAAATAGAA	5072
QY	5436	GTTA 5439	
Db	5073	GTTA 5076	
RESULT 4			
AF298548			
LOCUS	5100 bp	mRNA	linear PRI 19-FEB-2002
DEFINITION	Homo sapiens caspase recruitment domain protein 7 mRNA, complete cds.		
ACCESSION	AF298548		
VERSION	AF298548.1	GI:10198208	
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 5100)		
AUTHORS	Bertin,J. and DiStefano,P.S.		
TITLE	The PYRIN domain: a novel motif found in apoptosis and inflammation proteins		
JOURNAL	Cell Death Differ. 7 (12), 1273-1274 (2000)		
MEDLINE	21169419		
PUBMED	11270363		
REFERENCE	2 (bases 1 to 5100)		

AUTHORS	Bertin,J.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-AUG-2000) Neurobiology, Millennium Pharmaceuticals Inc., 640 Memorial Drive, Cambridge, MA 02139, USA		
FEATURES	Location/Qualifiers		
source	1..5100		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	160..4449		
CDS	/note="CARD7; nucleotide-binding site/leucine-rich repeat family member (NBS/LRR); contains PYRIN and CARD domains"		
	/codon_start=1		
	/product="caspase recruitment domain protein 7"		
	/protein_id="AAG15254.1"		
	/db_xref="GI:10198209"		
	/translation="MAGAWGRACLYLEFLKKEELKFPOLLANKANHSRSSGETPAQ PEKTSMEVSYLAQYGEQAWDLALHTWEQMLRSICAQAGAGHSPFPSPSE PHLGSPSTSTAVLPMWHELHPAGCTQGSERRVLRQLPTSGRRWREISASHLYQAL PSSPDHESPOESPNAFTSTAVLGSWSPQPSLAPRQEPAPGTQWPLDETSGLYYTE IREPEREKSEKRPMAAVCTPQAHSLSLOPHHPWEPSPRESLCSWPKNEDFNQ KFTQLLLQRPSPQDPLVRSNPDYVEENRHLIEIRDLEFGPLDQEPRIYVLQ AAGIKSTLARQVKEAMRGQLYGRFOHVFYSCRELAQSKVSLAELIGRDQATP APIRQLTSRPERLLFILDVDEPGWLVQEPSELCHWSQPQADALGSLGKTLIP BASFLITARTTALQNLIPSLQARWVEVLGFSESSRKEYFYRYFTDERQAIRAFRLV SNKELMALCLVPWSLACTCLMQMKRKBKLTLSKTTTTLCLHYLAQAQAOLPLG QLRDLCSLAESGIHQKTLFSPDLRKHGLDGAIIISTELKMGILQEHPIPLSYFIHL CFQFFAAMSVYLEDEKRGKHSNCIDLEKTLIAYGHLGFGASTTRFLGLSLDEG EREMENTPHCLSGQRNLMOWPSLQQLLQPHLSLESLHCLYETRNKTLPTQMAHFEE MCMVETDMELLVCTFCIKFSRHVKQLQIEGRQHSRTWSPMVVLFPMVPTDAYWQ ILFVLKVTNRNLKELDSGSLSAVSKLQRLRRPCLLETLRLVLAGCGLTAEDCKD LAFGLRNOTLITELDSLNVLMDSAGKLCORLQPSCKLQRLQSLGSLTSDCCQDL ASVLSASPSKELDLOQNLDDVGRLLCEGRHPACKLIRLGLDQTLTSDEMROELR ALBQKPLLIIFSRKESVMTPIEGLDGTGMSNTSSLKROPLRSERASAASHVAQNLK LLDVSKLPTPIAEIAEESPEVVPVELLVCVPSQSDHLHTKPLGTDGDFWGTGCVAT EVQDKENLYRHFVPAVSYPNPTGLCFVVRVAVTVEIECFVMDQFLGEINPQHSWM IVLENPSPIGLVLLKMHNLRIPTVTSVLLYRHLPEVTHLYLIPDSCSTRKE LELCVSPGEDQLSEFYVYDGLSGIRLOVKKDETLWEALVKRGLDMPATTLIPP ACTVSPFLDAPQLLHFVDYREGIIRKTSVEVVIDKLHQVLSQEQYERVLANTR PSQMRKLFSLSQSWDRCKDGLYQALKEHPHLMELMEKSKGLLPLSS"		
BASE COUNT	1251 a	1374 c	1406 g 1069 t
ORIGIN			
Query Match	89.7%	Score 4882.4;	DB 9; Length 5100;
Best Local Similarity	99.5%	Pred. No. 0;	
Matches 4898;	Conservative	0; Mismatches	26; Indels 0; Gaps 0;
QY	516	GACAGAGATGGCTGGCGGAGCCTGGGGCCGCTGGCCCTGTTACTTTGGAGTTCCTGAAGAA	575
Db	153	GNACAAAGATGGCTGGCGGAGCCTGGGGCCGCTGGCCCTGTTACTTTGGAGTTCCTGAAGAA	212
QY	576	GGAGGAGCTGAAGGATTCAGCTTCTGCTCGCCAAATAAAGCGCACTCCAGAGGCTTTC	635
Db	213	GGAGGAGCTGAAGGAGTTCAGCTTCTGCTCGCCAAATAAAGCGCACTCCAGAGGCTTTC	272
QY	636	GGGTGAGACACCCGCTCAGCCAGAGAAACGAGTGGCATGGAGTGGCGCTCGTACTCGT	695
Db	273	GGGTGAGACACCCGCTCAGCCAGAGAAACGAGTGGCATGGAGTGGCGCTCGTACTCGT	332
QY	696	GGCTCAGTATGGGAGCAGCGGGCCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGG	755
Db	333	GGCTCAGTATGGGAGCAGCGGGCCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGG	392
QY	756	GCTGAGGTCACTGTGCGCCCAAGCCAGGAAGGGGAGGCCACTCTCCCTCATTCCTCCCTA	815
Db	393	GCTGAGGTCACTGTGCGCCCAAGCCAGGAAGGGGAGGCCACTCTCCCTCATTCCTCCCTA	452
QY	816	CAGCCCAAGTGAACCCCACTGGGGTCTCCAGCCCAACCCACCTCCACCGAGTCTTAAT	875
Db	453	CAGCCCAAGTGAACCCCACTGGGGTCTCCAGCCCAACCCACCTCCACCGAGTCTTAAT	512
QY	876	GCCTTGGATCCATGAATTTGCCGGCGGGGTGCAACCAGGGCTCAGAGAGAAGGGTTTTGAG	935

Db 513 |||||GCCTGGATCATGAATTCGCGCGGGTGCACCCAGGGCTCAGAGAGAAGGGTTTGTAG 572
QY 936 ACAGCTGCTGTGACACATCTGGAGCGCGCTGGAGAGAAATCTCTGCTCTCACTCTCTPACCA 995
Db 573 ACAGCTGCTGTGACACATCTGGAGCGCGCTGGAGAGAAATCTCTGCTCTCACTCTPACCA 632
QY 996 AGCTCTTCCAAGCTCCCGACACATAGTCTCCAGCCAGGAGTCAACCAACGCCCCCAC 1055
Db 633 AGCTCTTCCNAGCTCCCGACACATAGTCTCCNAGCCAGGAGTCAACCAACGCCCCCAC 692
QY 1056 ATCCACAGAGTGTGGGAGCTGGGGATCCCACTCAGCCAGCGCTAGCACCCAGAGA 1115
Db 693 ATCCACAGAGTGTGGGAGCTGGGGATCCCACTCAGCCAGCGCTAGCACCCAGAGA 752
QY 1116 GCAGGAGGCTCTGGGACCCNATGGGCTCTGGATGAAAAGTCAAGAAATTTACTACACAGA 1175
Db 753 GCAGGAGGCTCTGGGACCCNATGGGCTCTGGATGAAAAGTCAAGAAATTTACTACACAGA 812
QY 1176 AATCAGAGAAAGAGAGAGAGAGAAATCAGAGAAAGGAGCGCCCAATGGGAGCGGTGGT 1235
Db 813 AATCAGAGAAAGAGAGAGAGAGAAATCAGAGAAAGGAGCGCCCAATGGGAGCGGTGGT 872
QY 1236 AGGAAACGCCCCACAGCGGACACAGAGCTACAGCCCCACAGCCACCCATGGGAGCGCTTC 1295
Db 873 AGGAAACGCCCCACAGCGGACACAGAGCTACAGCCCCACAGCCACCCATGGGAGCGCTTC 932
QY 1296 TGTGAGAGAGCGCTGTCTTCCATGCGCCCTGGGAAATAGGATTTTAAACCAAAATTT 1355
Db 933 TGTGAGAGAGCGCTGTCTTCCATGCGCCCTGGGAAATAGGATTTTAAACCAAAATTT 992
QY 1356 CACACAGCTGTACTTCTACAAAGACCTCACCCAGAGCAAGATCCCTGTCAAGAG 1415
Db 993 CACACAGCTGTACTTCTACAAAGACCTCACCCAGAGATCCCTGTCAAGAG 1052
QY 1416 AAGCTGCTGTGATTAATGTGGAGAGAAATCGAGGACATTTAAATGAGATCAGAGACTTAT 1475
Db 1053 AAGCTGCTGTGATTAATGTGGAGAGAAATCGAGGACATTTAAATGAGATCAGAGACTTAT 1112
QY 1476 TGGCCAGAGCGCTGGATACCCAGAACTCTGCTAGTATCATCTGCGGGGCTGCTGGAT 1535
Db 1113 TGGCCAGAGCGCTGGATACCCAGAACTCTGCTAGTATCATCTGCGGGGCTGCTGGAT 1172
QY 1536 TGGGAAAGTCAACACTGCGCCAGGAGGTGAAGGAAGCTGGGGAGAGGCCAGCTGTATGG 1595
Db 1173 TGGGAAAGTCAACACTGCGCCAGGAGGTGAAGGAAGCTGGGGAGAGGCCAGCTGTATGG 1232
QY 1596 GGAACGCTTCCAGCATGTCTTCTACTTTCAGCTCGAGAGAGTGGCCCAAGTCCAAAGTGGT 1655
Db 1233 GGAACGCTTCCAGCATGTCTTCTACTTTCAGCTCGAGAGAGTGGCCCAAGTCCAAAGTGGT 1292
QY 1656 GAGTCTCGCTGAGCTCATCGGAAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGAT 1715
Db 1293 GAGTCTCGCTGAGCTCATCGGAAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGAT 1352
QY 1716 CCGTCTTAGGCCAGAGCGGCTGCTTCTATCTCGATGTGTAGATGAGCAGAGATGGGT 1775
Db 1353 CCGTCTTAGGCCAGAGCGGCTGCTTCTATCTCGATGTGTAGATGAGCAGAGATGGGT 1412
QY 1776 CTTTGCAGGAGCCGAGTTTCTGAGCTCTGTCTGCACTGGAGCCAGCCAGCGCGGATGC 1835
Db 1413 CTTTGCAGGAGCCGAGTTTCTGAGCTCTGTCTGCACTGGAGCCAGCCAGCGCGGATGC 1472
QY 1836 ACTGTGGGAGGTTTGTCTGGGAAAACTATATCTCCGAGGCACTCTTCTGTATCAGGC 1895
Db 1473 ACTGTGGGAGGTTTGTCTGGGAAAACTATATCTCCGAGGCACTCTTCTGTATCAGGC 1532
QY 1896 TCGGACACAGCTCTGCAGAACTCATCTTCTCTTTGGAGCAGCGTGGGTAGAGT 1955
Db 1533 TCGGACACAGCTCTGCAGAACTCATCTTCTCTTTGGAGCAGCGTGGGTAGAGT 1592
QY 1956 CTTGGGGTCTCTGAGTCCAGCAGGAGGAATTTTCTACAGATATTTTACAGATGAAAG 2015

Db 1593 CTTGGGGTCTCTGAGTCCAGCAGGAAGAAATATTTTCTACAGATATTTTACAGATGAAG 1652
QY 2016 GCAACAAATTAGAGCCTTTTAGGTTGGTCAAAATCAAAACAAAGAGCTCTGGGCGCTGTGTCT 2075
Db 1653 GCNAGCAATTAGAGCCTTTTAGGTTGGTCAAAATCAAAACAAAGAGCTCTGGGCGCTGTGTCT 1712
QY 2076 TGTGCGCTGGGTGTCTGGCTGGCTGCACTTTGCTGTGATGAGCAGAGATGAACGGGAAGA 2135
Db 1713 TGTGCGCTGGGTGTCTGGCTGGCTGCACTTTGCTGTGATGAGCAGAGATGAACGGGAAGA 1772
QY 2136 AAAATCTACACTGACTCTCCNAGACCAACCAACCCCTCTGTCTTACATTAACCTTGGCCAGGC 2195
Db 1773 AAAATCTACACTGACTCTCCNAGACCAACCAACCCCTCTGTCTTACATTAACCTTGGCCAGGC 1832
QY 2196 TCTCAAGCTCAGCCATTTGGGACCCAGCTCAGAGACCTCTCTCTCTGGCTGTGAGGG 2255
Db 1833 TCTCCNAGCTCAGCCATTTGGGACCCAGCTCAGAGACCTCTCTCTCTGGCTGTGAGGG 1892
QY 2256 CATCTGGCAAAAAAGACCCCTTTTTCAGTCCAGATGACCTCAGGAAAGCATGGGTATGATGG 2315
Db 1893 CATCTGGCAAAAAAGACCCCTTTTTCAGTCCAGATGACCTCAGGAAAGCATGGGTATGATGG 1952
QY 2316 GGCCTCATCTCCACCTTCTTGAAGATGGTATCTTCAAGAGCACCCCATCCCTCTGAG 2375
Db 1953 GGCCTCATCTCCACCTTCTTGAAGATGGTATCTTCAAGAGCACCCCATCCCTCTGAG 2012
QY 2376 CTACAGCTTCATTCACCTCTGTTTCAAGAGTCTTTTGCAGCAATGTCTATGTCTTGA 2435
Db 2013 CTACAGCTTCATTCACCTCTGTTTTCAGGAGTCTTTGCAGCAATGTCTATGTCTTGA 2072
QY 2436 GGATCAGAAAGGGGAGAGGTAAACATTTCTAATTTGCATCATAGATTTTGGAAAAAGCGCTAGA 2495
Db 2073 GGATCAGAAAGGGGAGAGGTAAACATTTCTAATTTGCATCATAGATTTTGGAAAAAGCGCTAGA 2132
QY 2496 AGCATATGGAATACATGGCTGTTTGGGCGATCAACCAACGTTTCTTATTTGGGCGCTGT 2555
Db 2133 AGCATATGGAATACATGGCTGTTTGGGCGATCAACCAACGTTTCTTATTTGGGCGCTGT 2192
QY 2556 AAGTATGAGGGGAGAGAGATGGAGAAACATCTTTTCACTGCGGCTCTCTCAGGGAG 2615
Db 2193 AAGTATGAGGGGAGAGAGATGGAGAAACATCTTTTCACTGCGGCTCTCTCAGGGAG 2252
QY 2616 GAACTGTATGCAAGTGGGTCCGCTCCAGCTGTGTGTCAGGCCACACTCTCTGAGGTC 2675
Db 2253 GAACTGTATGCAAGTGGGTCCGCTCCAGCTGTGTGTCAGGCCACACTCTCTGAGGTC 2312
QY 2676 CTTCACTGCTTTGTACGAGACTCGGAAACAAACGTTTCTTGACACAGTCAATGGCCCATTT 2735
Db 2313 CTTCACTGCTTTGTATGAGACTCGGAAACAAACGTTTCTTGACACAGTCAATGGCCCATTT 2372
QY 2736 CGAAGAAATGGGCAATGTGTGTAGAAACAGACATGGAGCTCTTAGTGTGCATTTCTGCAT 2795
Db 2373 CGAAGAAATGGGCAATGTGTGTAGAAACAGACATGGAGCTCTTAGTGTGCATTTCTGCAT 2432
QY 2796 TAAATTCAGCCCGCAGCTGAAGAAAGCTTCAGCTGATTTGAGGGCAGGCAGCAGATCAAC 2855
Db 2433 TAAATTCAGCCCGCAGCTGAAGAAAGCTTCAGCTGATTTGAGGGCAGGCAGCAGATCAAC 2492
QY 2856 ATGGAGCCCAACATGGTGTCTGTTTCAAGTGGGTCCAGTCAACAGATGCTTATTTGGCA 2915
Db 2493 ATGGAGCCCAACATGGTGTCTGTTTCAAGTGGGTCCAGTCAACAGATGCTTATTTGGCA 2552
QY 2916 GATTCCTTCTCGGCTCAAGGTCACAGAAACCTGAGGAGCTGGACCTAAGTGGAA 2975
Db 2553 GATTCCTTCTCGGCTCAAGGTCACAGAAACCTGAGGAGCTGGACCTAAGTGGAA 2612
QY 2976 CTTGCTGAGCCCACTGTGCAAGTGAAGTCTTTTGAAGACCTCTGAGACGCCCTCGCTGCCT 3035
Db 2613 CTTGCTGAGCCCACTGTGCAAGTGAAGTCTTTTGAAGACCTCTGAGACGCCCTCGCTGCCT 2672
QY 3036 CTTGAGACCCCTGCGGTTGGCTGTGGCTCTCAAGCTGAGGACTGCAAGAGACCTTGC 3095
Db 2673 CTTGAGACCCCTGCGGTTGGCTGTGGCTCTCAAGCTGAGGACTGCAAGAGACCTTGC 2732

QY 3096 CTTTGGGCTGAGAGCCAAACAGACCCTGACCGAGCTGGACCTGAGCTTCAATGTGTCTAC 3155
 Db 2733 CTTTGGGCTGAGAGCCAAACAGACCCTGACCGAGCTGGACCTGAGCTTCAATGTGTCTAT 2792
 QY 3156 GGATGCTGAGAGCCAAACAGCTTTGGCAGAGACTGAGACAGCGGAGGTGCAAGCTACAGCG 3215
 Db 2793 GGATGCTGAGAGCCAAACAGCTTTGGCAGAGACTGAGACAGCGGAGGTGCAAGCTACAGCG 2852
 QY 3216 ACTGAGCTGGTCACTGTGGCTCAGCTCAGCTGACTGCTGCCAGGACCTGGCCTCTGTGCT 3275
 Db 2853 ACTGAGCTGGTCACTGTGGCTCAGCTCAGCTGCTGCCAGGACCTGGCCTCTGTGCT 2912
 QY 3276 TAGTGCCAGCCCGAGCTGAAAGAGCTAGACCTGACAGAGAAACAACTCGATGACGTTGG 3335
 Db 2913 TAGTGCCAGCCCGAGCTGAAAGAGCTAGACCTGACAGAGAAACAACTCGATGACGTTGG 2972
 QY 3336 CBTGCGACTGCTGTGAGGGCTCAGGCACTCTGCTGCTGCAAACTATACGCTGGGGCT 3395
 Db 2973 CBTGCGACTGCTGTGAGGGCTCAGGCACTCTGCTGCTGCAAACTATACGCTGGGGCT 3032
 QY 3396 GGACAGAGCAACTCTGAGTGATGAGATGAGGAGGAGCTGAGGGCCCTGGAGCAGAGAA 3455
 Db 3033 GGACAGAGCAACTCTGAGTGATGAGATGAGGAGGAGCTGAGGGCCCTGGAGCAGAGAA 3092
 QY 3456 ACTCAGCTGCTCATCTTCAGCAGACCGAAACCAAGTGTGATGACCCCTACTGAGGGCT 3515
 Db 3093 GCTCAGCTGCTCATCTTCAGCAGACCGAAACCAAGTGTGATGACCCCTACTGAGGGCT 3152
 QY 3516 GGATCGGAGAGATGAGTAATAGCACTCTCACTCAAGCGGCGAGACTCGGATCAGA 3575
 Db 3153 GGATCGGAGAGATGAGTAATAGCACTCTCACTCAAGCGGCGAGACTCGGATCAGA 3212
 QY 3576 GAGGGCGCTTCCCATGTTGCTCAGGCTAATCTCAAACTCTCGGAGCTGAGCAAGATCTT 3635
 Db 3213 GAGGGCGCTTCCCATGTTGCTCAGGCTAATCTCAAACTCTCGGAGCTGAGCAAGATCTT 3272
 QY 3636 CCCAATTGCTGAGATTGACAGAGAAAGCTCCCGAGAGTGTACCGGTGGAACCTTTGTG 3695
 Db 3273 CCCAATTGCTGAGATTGACAGAGAAAGCTCCCGAGAGTGTACCGGTGGAACCTTTGTG 3332
 QY 3696 CBTGCTTCTCTGCTCTCAAGGGGACCTGCATACGAAGCCCTTTGGGACGTGACGATGA 3755
 Db 3333 CBTGCTTCTCTGCTCTCAAGGGGACCTGCATACGAAGCCCTTTGGGACGTGACGATGA 3392
 QY 3756 CTTCTGGGGCCCAACGGGGCTGTGGCTACTGAGTGTGTAAGAAAGAACTTTGTA 3815
 Db 3393 CTTCTGGGGCCCAACGGGGCTGTGGCTACTGAGGTAGTGTGAAGAAAGAACTTTGTA 3452
 QY 3816 CCGAGTTCACTTCCCTGTAGCTGGCTCTACCGCTGGCCCAACACGGGTCTCTGCTTTGT 3875
 Db 3453 CCGAGTTCACTTCCCTGTAGCTGGCTCTACCGCTGGCCCAACACGGGTCTCTGCTTTGT 3512
 QY 3876 GATGAGAGAGCGGTGACCGTTGAGATTGAAATTTGTGTGTGGGACCACTTCTGGGTGA 3935
 Db 3513 GGTGAGAGAGCGGTGACCGTTGAGATTGAAATTTGTGTGTGGGACCACTTCTGGGTGA 3572
 QY 3936 GATCAACCCACAGCAGCTGATGTGGCAGGGCCCTCTGTGGACATCAAGGCTGAGCC 3995
 Db 3573 GATCAACCCACAGCAGCTGATGTGGCAGGGCCCTCTGTGGACATCAAGGCTGAGCC 3632
 QY 3996 TGGAGCTGTGGAAGCTGTGCACCTCCCTCACTTTGTGGCTCTCCAGGGGGCCATGTGGA 4055
 Db 3633 TGGAGCCGTGTGGAAGCTGTGCACCTCCCTCACTTTGTGGCTCTCCAGGGGGCCATGTGGA 3692
 QY 4056 CACATCCTGTTCCAAATGGCCCACTTTAAAGAGGAGGGGATGCTCTCTGGAGAGCCAGC 4115
 Db 3693 CACATCCTGTTCCAAATGGCCCACTTTAAAGAGGAGGGGATGCTCTCTGGAGAGCCAGC 3752
 QY 4116 CAGGCTGAGCTGTCATCAGATAGTCTTGAAACCCAGCTTCTCCCCCTTGGGAGTCT 4175
 Db 3753 CAGGCTGAGCTGTCATCAGATAGTCTTGAAACCCAGCTTCTCCCCCTTGGGAGTCT 3812

QY 4176 CTTGAAAAATGATCCATAATGCCCTGCGCTTCATTCCCTCACCCTCTGTGGTGTCTTTA 4235
 Db 3813 CTTGAAAAATGATCCATAATGCCCTGCGCTTCATTCCCTCACCCTCTGTGGTGTCTTTA 3872
 QY 4236 CCACCGCTTCATCTCTGAGGAGTCACTTCCACTCTACCTCTACCTGATCCCAAGTACTGTC 4295
 Db 3873 CCACCGCTTCATCTCTGAGGAGTCACTTCCACTCTACCTCTACCTGATCCCAAGTACTGTC 3932
 QY 4296 CATTCGGAAGAACTGGAGCTCTGCTATCGAAGCCCTCGAAGAACACCAAGCTGTTCTCGGA 4355
 Db 3933 CATTCGGAAGAACTGGAGCTCTGCTATCGAAGCCCTCGAAGAACACCAAGCTGTTCTCGGA 3992
 QY 4356 GTTCTACCTGTGGCCACTTTGGGATCAGGATCAGGCTGCAAGTGAAGAACCAAGATGA 4415
 Db 3993 GTTCTACCTGTGGCCACTTTGGGATCAGGATCAGGCTGCAAGTGAAGAACCAAGATGA 4052
 QY 4416 GACTCTGTGTGGAGGGCTTTGGTGAACCAAGGAGATCTCATGCTGCAACTACTCTCAT 4475
 Db 4053 GACTCTGTGTGGAGGGCTTTGGTGAACCAAGGAGATCTCATGCTGCAACTACTCTCAT 4112
 QY 4476 CCCTCCAGCCCGCATAGCCGTACCTTTCACCTCTGATGCCCGCAGTTGCTGCATTTGT 4535
 Db 4113 CCCTCCAGCCCGCATAGCCGTACCTTTCACCTCTGATGCCCGCAGTTGCTGCATTTGT 4172
 QY 4536 GGACAGTATCGAGAGCAGCTGATAGCCCGAGTGCATCGGTGGAGGTTGTCTTGGACAA 4595
 Db 4173 GGACAGTATCGAGAGCAGCTGATAGCCCGAGTGCATCGGTGGAGGTTGTCTTGGACAA 4232
 QY 4596 ACTGATGGAAGCTGTGAGCCAGGAGCAGTACAGAGGGTGTGGCTGAGAACACAG 4655
 Db 4233 ACTGATGGAAGCTGTGAGCCAGGAGCAGTACAGAGGGTGTGGCTGAGAACACAG 4292
 QY 4656 GCCAGCCAGATGCGGAAGCTGTTTCAGCTTGAGCCAGTCTCGGACCGGAAGTGCAAGA 4715
 Db 4293 GCCAGCCAGATGCGGAAGCTGTTTCAGCTTGAGCCAGTCTCGGACCGGAAGTGCAAGA 4352
 QY 4716 TGGACTCTACCAAGCCCTGAAAGGAGACCCATCTCCTCATTTATGGAACCTGSGAGAA 4775
 Db 4353 TGGACTCTACCAAGCCCTGAAAGGAGACCCATCTCCTCATTTATGGAACCTGSGAGAA 4412
 QY 4776 GGGCAGAAAAAGGAGCTCTTGCACCTCAGCAGCTGAAGTATCAACACACAGCCCTTGACC 4835
 Db 4413 GGGCAGAAAAAGGAGCTCTTGCACCTCAGCAGCTGAAGTATCAACACACAGCCCTTGACC 4472
 QY 4836 CTTGAGCTCTGCTTTGCTGACCTCTTTGGGTCTCAGTTTCTCTCTGCAAAACAG 4895
 Db 4473 CTTGAGCTCTGCTTTGCTGACCTCTCTTTGGGTCTCAGTTTCTCTCTGCAAAACAG 4532
 QY 4896 TTGCCATCTGCTTTGCCCTTCCAGCAGCTAAAGTAAATGGAACCTTTGATGATGCTTTGCTGG 4955
 Db 4533 TTGCCATCTGCTTTGCCCTTCCAGCAGCTAAAGTAAATGGAACCTTTGATGATGCTTTGCTGG 4592
 QY 4956 GCAATTATGTCATGTCAGGAGTGCACAGGGGGCCCGAGTCCAGGTGGCCCTAAACAG 5015
 Db 4593 GCAATTATGTCATGTCAGGAGTGCACAGGGGGCCCGAGTCCAGGTGGCCCTAAACAG 4652
 QY 5016 TCTCAGGGAATGTCCATCTGAGCTGCAAGCCCTCGACAGCTCATAGAGCTCATCT 5075
 Db 4653 TCTCAGGGAATGTCCATCTGAGCTGCAAGCCCTCGACAGCTCATAGAGCTCATCT 4712
 QY 5076 GTTGCCACAGCAGCCCAAGCCCTAGAGCCCTCCGATCCATCCAGGGCCCAAGAGGAATA 5135
 Db 4713 GTTGCCACAGCAGCCCAAGCCCTAGAGCCCTCCGATCCATCCAGGGCCCAAGAGGAATA 4772
 QY 5136 GGAGGGAATGGAACCAATTTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5195
 Db 4773 GGAGGGAATGGAACCAATTTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4832
 QY 5196 TCAGGCTGGGAGGGCCAGCTGT 5255
 Db 4833 TCAGGCTGGGAGGGCCAGCTGT 4892
 QY 5256 AACAGAGTAAAGTGGAAGGAGTTTATTTCAGAAAAATAAGAGGATATCAGCTCTTTTAG 5315

Db	1501	TTGGGAAGTCAA	CAGTGGCAGG	TGAAGGAGCCT	GGGAGAGGCC	AGCTGTATG	1560
Qy	1595	GGGACCGCTTCAG	CAATGCTTCTACT	CAGCTGCAGAG	AGCTGGGCCAGT	CCAAGGTGG	1654
Db	1561	GGGACCGCTTCC	AGCATGTCTTCT	TACTT	CAGCTGCAGAG	AGCTGGGCCAGT	1620
Qy	1655	TGAGTCTCGCTG	AGCTCATCGG	AAAGATGGG	ACGCCACTCGG	CTCCATT	1714
Db	1621	TGAGTCTCGCTG	AGCTCATCGG	AAAGATGGG	ACGCCACTCGG	CTCCATT	1680
Qy	1715	TCCTGTCTAGGC	CAGACGGCTGCT	CTT	CATCTCGAT	GTGTAGATGAG	1774
Db	1681	TCCTGTCTAGGC	CAGACGGCTGCT	CTT	CATCTCGAT	GTGTAGATGAG	1740
Qy	1775	TCCTGAGAGCGG	AGTCTGAGCTCT	GTCTGTG	ACCTGAGCCAC	AGCCGCGGGATG	1834
Db	1741	TCCTGAGAGCGG	AGTCTGAGCTCT	GTCTGTG	ACCTGAGCCAC	AGCCGCGGGATG	1800
Qy	1835	CAGTCTGGGCGA	GTGCTGGGGAAC	TACTT	CCCGAGGCAT	CCTTCTGAT	1894
Db	1801	CAGTCTGGGCGA	GTGCTGGGGAAC	TACTT	CCCGAGGCAT	CCTTCTGAT	1860
Qy	1895	CTCGACACAGCT	CTGCGAAGCCT	CA	TCTTCTT	TGGAGCAGC	1954
Db	1861	CTCGACACAGCT	CTGCGAAGCCT	CA	TCTTCTT	TGGAGCAGC	1920
Qy	1955	TCCTGGGGTCT	CTGAGTCAGC	AGGAAGATAT	TCTT	TACAGATAT	2014
Db	1921	TCCTGGGGTCT	CTGAGTCAGC	AGGAAGATAT	TCTT	TACAGATAT	1980
Qy	2015	GGCAAGCAATTA	GAGCTTTAGT	GGTCAAA	TCAAA	CAGAGCTCT	2074
Db	1981	GGCAAGCAATTA	GAGCTTTAGT	GGTCAAA	TCAAA	CAGAGCTCT	2040
Qy	2075	TTGTGCGCTGG	GTGCTGGCTG	CGCTGC	ATTTGCTGAT	GCAGAGATGA	2134
Db	2041	TTGTGCGCTGG	GTGCTGGCTG	CGCTGC	ATTTGCTGAT	GCAGAGATGA	2100
Qy	2135	AAAACTCACTGA	CTTCAAGACAC	CAACAC	CCCTCTGT	CTACATTA	2194
Db	2101	AAAACTCACTGA	CTTCAAGACAC	CAACAC	CCCTCTGT	CTACATTA	2160
Qy	2195	CTCTCAAGCTC	AGCCATTTGGG	ACCCAGCT	CTGCTG	CTGCTGCTG	2254
Db	2161	CTCTCAAGCTC	AGCCATTTGGG	ACCCAGCT	CTGCTG	CTGCTGCTG	2220
Qy	2255	GCATCTGGCA	AAAAAGACCC	TTTT	CAGTCC	AGATGAGTGG	2314
Db	2221	GCATCTGGCA	AAAAAGACCC	TTTT	CAGTCC	AGATGAGTGG	2280
Qy	2315	GGGCCATCAT	CTCCACCTT	TTGAA	AGATGGGAT	TTCTT	2374
Db	2281	GGGCCATCAT	CTCCACCTT	TTGAA	AGATGGGAT	TTCTT	2340
Qy	2375	GCTACAGCTTCA	TACCTCTGTTT	CCAGAGT	TTCTT	CCAGCAATG	2434
Db	2341	GCTACAGCTTCA	TACCTCTGTTT	CCAGAGT	TTCTT	CCAGCAATG	2400
Qy	2435	AGGATGAGA	AGGGAGAGGT	TAACAT	TTCTA	ATTTGCAAT	2494
Db	2401	AGGATGAGA	AGGGAGAGGT	TAACAT	TTCTA	ATTTGCAAT	2460
Qy	2495	AAGCATATGGA	ATATAGCGCT	GTGTTTGGG	GCATCAAC	CAAGCTTT	2554
Db	2461	AAGCATATGGA	ATATAGCGCT	GTGTTTGGG	GCATCAAC	CAAGCTTT	2520
Qy	2555	TAAGTATGAG	GGGAGAGAGAT	GGAGACAT	TTCTT	CACTGCGGCT	2614
Db	2521	TAAGTATGAG	GGGAGAGAGAT	GGAGACAT	TTCTT	CACTGCGGCT	2580
Qy	2615	GGAACTGATGA	GTGGGTCCCGT	CTCCG	CTGCTG	CAGCCACACT	2674

Db	2581	GGAACTGATGA	GTGGGTCCCGT	CTCCG	CTGCTG	CAGCCACACT	2640
Qy	2675	CCCTCCACTGT	TTGTAACAGAG	CTCGG	AAACAA	CTTCTGAC	2734
Db	2641	CCCTCCACTGT	TTGTAACAGAG	CTCGG	AAACAA	CTTCTGAC	2700
Qy	2735	TCGAAAGAAAT	GGGCATGTGT	GTAGAAA	CAGACAT	GGAGCTT	2794
Db	2701	TCGAAAGAAAT	GGGCATGTGT	GTAGAAA	CAGACAT	GGAGCTT	2760
Qy	2795	TTAAATTCAGC	CGCCACAGTGA	AGAGCTT	CAAGTGA	GGGAGGAG	2854
Db	2761	TTAAATTCAGC	CGCCACAGTGA	AGAGCTT	CAAGTGA	GGGAGGAG	2820
Qy	2855	CATGGAGCC	CCACCATG	CTGAGT	CTGTT	CAGTGGT	2914
Db	2821	CATGGAGCC	CCACCATG	CTGAGT	CTGTT	CAGTGGT	2880
Qy	2915	AGATTCTCT	CTTCTCCG	CTCCTCA	AGGTCAC	AGAACTT	2974
Db	2881	AGATTCTCT	CTTCTCCG	CTCCTCA	AGGTCAC	AGAACTT	2940
Qy	2975	ACTCGCTGAG	CCACTCTG	CGAGTGA	AGTCTT	TTGTAAC	3034
Db	2941	ACTCGCTGAG	CCACTCTG	CGAGTGA	AGTCTT	TTGTAAC	3000
Qy	3035	TCCTGGAG	ACCTCGG	TTGGCTG	CTGCTG	CCACAGCT	3094
Db	3001	TCCTGGAG	ACCTCGG	TTGGCTG	CTGCTG	CCACAGCT	3060
Qy	3095	CTTTGGGCTG	AGAGCCCA	ACAGAC	CTTCA	CGAGCTG	3154
Db	3061	CTTTGGGCTG	AGAGCCCA	ACAGAC	CTTCA	CGAGCTG	3120
Qy	3155	CGGATGCTG	GGAGCCAAA	CACCTT	TTGCC	AGAGACT	3214
Db	3121	CGGATGCTG	GGAGCCAAA	CACCTT	TTGCC	AGAGACT	3180
Qy	3215	GACTGAC	CTGGT	CTGCTG	CTGCTG	CTGCTG	3274
Db	3181	GACTGAC	CTGGT	CTGCTG	CTGCTG	CTGCTG	3240
Qy	3275	TTAGTCC	CAGCCCA	CGAGCTG	AGAGCT	CTGAGC	3334
Db	3241	TTAGTCC	CAGCCCA	CGAGCTG	AGAGCT	CTGAGC	3300
Qy	3335	GGTGGCA	CTGCTG	TGAGGGG	CTCAG	GCATCT	3394
Db	3301	GGTGGCA	CTGCTG	TGAGGGG	CTCAG	GCATCT	3360
Qy	3395	TGGAC	CAGAC	CACTCTG	AGTGA	TGAGGAG	3454
Db	3361	TGGAC	CAGAC	CACTCTG	AGTGA	TGAGGAG	3420
Qy	3455	AACCTC	CAGCTG	CTCAT	TTT	CAGCAG	3514
Db	3421	AACCTC	CAGCTG	CTCAT	TTT	CAGCAG	3480
Qy	3515	TGGATAC	CGGAG	AGATG	ATATAG	CACTCT	3574
Db	3481	TGGATAC	CGGAG	AGATG	ATATAG	CACTCT	3540
Qy	3575	AGAGG	CGGCTT	CCCAT	TTGCT	CAGGCT	3634
Db	3541	AGAGG	CGGCTT	CCCAT	TTGCT	CAGGCT	3600
Qy	3635	TCCCAAT	TGCTG	AGATT	TCAG	AGAACT	3694
Db	3601	TCCCAAT	TGCTG	AGATT	TCAG	AGAACT	3660
Qy	3695	GGTGC	CTTCT	CTGCT	CTT	CAAGGG	3754
Db	3661	GGTGC	CTTCT	CTGCT	CTT	CAAGGG	3720

QY 3755 ACTTCTGGGGCCCAACGGGGCCCTGTGGCTACTGAGGTAGTTGACAAAGAAAAGAACTTCT 3814
Db 3721 ACTTCTGGGGCCCAACGGGGCCCTGTGGCTACTGAGGTAGTTGACAAAGAAAAGAACTTCT 3780
QY 3815 ACCGAGTTCATCTTCCCTGTAGCTGGCTCCTACCGCTGGCCCAACACGGGTCTCTGCTTTG 3874
Db 3781 ACCGAGTTCATCTTCCCTGTAGCTGGCTCCTACCGCTGGCCCAACACGGGTCTCTGCTTTG 3840
QY 3875 TGATGAGAGAAGCGGTGACCGTTGAGATTGAATTCCTGTGTGGGACCAAGTTCTCTGGGTG 3934
Db 3841 TGATGAGAGAAGCGGTGACCGTTGAGATTGAATTCCTGTGTGGGACCAAGTTCTCTGGGTG 3900
QY 3935 AGATCAACCCACAGCACAGCTGGATGGTGGCAGGGCTCTCTGTGGACATCAAGGCTGAGC 3994
Db 3901 AGATCAACCCACAGCACAGCTGGATGGTGGCAGGGCTCTCTGTGGACATCAAGGCTGAGC 3960
QY 3995 CTGGAGCTGTGGAAGCTGTGCACCTCCCTCACTTTGTGGCTCTCAAGGGGGCCATGTGG 4054
Db 3961 CTGGAGCTGTGGAAGCTGTGCACCTCCCTCACTTTGTGGCTCTCAAGGGGGCCATGTGG 4020
QY 4055 ACACATCCCTGTTCCAAATGGCCCACTTTAAAGAGGAGGGGATGCTCTCTGGAGAAGCCAG 4114
Db 4021 ACACATCCCTGTTCCAAATGGCCCACTTTAAAGAGGAGGGGATGCTCTCTGGAGAAGCCAG 4080
QY 4115 CCAGGCTGGAGCTGCATCATAGTTCTGGAAAACCCCACTTCTCCCTCTGGAGTCC 4174
Db 4081 CCAGGCTGGAGCTGCATCATAGTTCTGGAAAACCCCACTTCTCCCTCTGGAGTCC 4140
QY 4175 TCCTGAAATGATCCATAATGCTCGGCTTCATTCCTCGCTACCTCTGTGGTGTGCTTT 4234
Db 4141 TCCTGAAATGATCCATAATGCTCGGCTTCATTCCTCGCTACCTCTGTGGTGTGCTTT 4200
QY 4235 ACCACGGCTCCATCTCTGAGGAGTCACTTCCACCTCTACCTGATCCCAAGTACTGCT 4294
Db 4201 ACCACGGCTCCATCTCTGAGGAGTCACTTCCACCTCTACCTGATCCCAAGTACTGCT 4260
QY 4295 CCATTCGG----- 4302
Db 4261 CCATTCGGAAGGCCATAGATGATCTAGAAATGAATTCAGTTGTGCGAATCCACAAGC 4320
QY 4303 ----- 4302
Db 4321 CACCCCGCTGACCCCACTTTATATGGGCTGTCTTACACTGTCTGGTCTGTTTCAAG 4380
QY 4303 -----AAGGAAGTGGAGCTCTGCTATCGAAGCCCTGGAGAGACC 4342
Db 4381 GGATGTGGAATACTCCCAAGGAAGTGGAGCTCTGCTATCGAAGCCCTGGAGAGACC 4440
QY 4343 AGCTGTCTCGGAGTCTACGTTGGCCACTTGGGATCAGGGATCAGGCTGCAAGTGAAAG 4402
Db 4441 AGCTGTCTCGGAGTCTACGTTGGCCACTTGGGATCAGGGATCAGGCTGCAAGTGAAAG 4500
QY 4403 ACAAGAAAGATGAGACTCTGGTGTGGAGCCCTTGTGTAACACAGAGATCTCATGCTTG 4462
Db 4501 ACAAGAAAGATGAGACTCTGGTGTGGAGCCCTTGTGTAACACAGAGATCTCATGCTTG 4560
QY 4463 CAACTACTCTGATCCCTCAGCCCGCATAGCCGTAACCTTCACTCTGTGATGCCCGCAGT 4522
Db 4561 CAACTACTCTGATCCCTCAGCCCGCATAGCCGTAACCTTCACTCTGTGATGCCCGCAGT 4620
QY 4523 TGCTGCACTTTGTGGACCAAGTATCGAGAGCAGCTGATAGCCCGAGTGACATCGGTGGAG 4582
Db 4621 TGCTGCACTTTGTGGACCAAGTATCGAGAGCAGCTGATAGCCCGAGTGACATCGGTGGAG 4680
QY 4583 TTGCTTTGGAACAACTGCATGGAAGTGTGTGAGCCAGGAGCAGTACGAGAGGGTCTGCG 4642
Db 4681 TTGCTTTGGAACAACTGCATGGAAGTGTGTGAGCCAGGAGCAGTACGAGAGGGTCTGCG 4740
QY 4643 CTGAGAACACGAGGCCCGCAGATGCGGAAGCTGTTAGCTTTGACGCCAGTCTCTGGAGC 4702
Db 4741 CTGAGAACACGAGGCCCGCAGATGCGGAAGCTGTTAGCTTTGACGCCAGTCTCTGGAGC 4800

QY 4703 GGAAGTGCAAGATGGAAGCTTACCAAGCCCTGAAAGGAGACCCATCTCCTACCTCATTTATGG 4762
Db 4801 GGAAGTGCAAGATGGAAGCTTACCAAGCCCTGAAAGGAGACCCATCTCCTACCTCATTTATGG 4860
QY 4763 AACTCTGGGAGAGAGGAGCAAAAGAGGACTCTCTCCACTCAGCAGCTGAACTATCAACA 4822
Db 4861 AACTCTGGGAGAGAGGAGCAAAAGAGGACTCTCTCCACTCAGCAGCTGAACTATCAACA 4920
QY 4823 CCAGCCCTTGACCCCTTGAGTCTCTGGCTTGGCTGACCCCTTCTTTGGGTCTCAGTTCTTT 4882
Db 4921 CCAGCCCTTGACCCCTTGAGTCTCTGGCTTGGCTGACCCCTTCTTTGGGTCTCAGTTCTTT 4980
QY 4883 CTCTGCAACAAAGTGGCCATCTGGTTTGGCTTCCAGCACTAAAGTAATGGAACCTTTGATG 4942
Db 4981 CTCTGCAACAAAGTGGCCATCTGGTTTGGCTTCCAGCACTAAAGTAATGGAACCTTTGATG 5040
QY 4943 ATGCTCTTCTGGGCAATATGTCATGCCAGGATGCCACAGGGGGCC 4992
Db 5041 GATGCTTCTGGGCAATATGTCATGCCAGGATGCCACAGGGGGGGCC 5089

RESULT 6
AP229059
LOCUS
DEFINITION
Homo sapiens NAC-alpha splice variant (NAC) mRNA, complete cds,
alternatively spliced.
ACCESSION
AP229059
VERSION
AP229059.1 GI:12656104
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 4290)
AUTHORS
Chu,Z.-L., Pio,F., Xie,Z., Godzik,A. and Reed,J.C.
TITLE
NAC: an Apaf-1/Ced-4 family member regulates the cytochrome c
pathway for apoptosis
JOURNAL
Unpublished
2 (bases 1 to 4290)
AUTHORS
Chu,Z.-L. and Reed,J.C.
TITLE
Direct Submission
JOURNAL
Submitted (28-JAN-2000) Apoptosis and Signal Transduction, The
Burnham Institute, 10901 North Torrey Pines Road, La Jolla, CA
92037, USA

FEATURES
Location/Qualifiers
1..4290
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="11"
/cell_type="activated Jurkat T cells"
1..4290
/gene="NAC"
1..4290
/gene="NAC"
/note="Apaf-1/Ced-4 family apoptosis regulator; contains
NB-ABC domain, CARD domain, and leucine-rich repeats
(LRR)"
/codon_start=1
/product="NAC-alpha splice variant"
/protein_id="AAK00748.1"
/db_xref="GI:12656105"
/translation="NAGAWGRACLYLFLKKEELKEFOLLILANKAHSSSGETPAQ
PEKTSQMVASYLVAYQYGEQRANDLALHTWEQMRSLCAQAEAGAGHSFSPYFSE
PHLGSPTSTAVLPMWHELPACTQGSERRVLRQPDTSGRWRREISALSLYAL
PSPSPHSPSESNAPTSTAVLGSWSPQPSLAPREQEPGTQWPLDETSGIYYTE
IRERREKSEKRPWPAVVGTPQAAHSLQPHHPWSPVRESLCSLTPWKNDFNQ
KFTOLLQRPHPRSQDPLVKSPDYVEENRHLIEIRDLFGPLDTPRIVILQG
RAGIKSTLARGVKEAWRGOLYGRFOHVFESCRELAQSKVSLAEIIGDGTATP
APIQLRSPERLLFILGDVFGWVQEPSELCLHWSQPPADALGSLGKTLIP
EASFLITATLQNLIPSEQRVWVLFVGSSESKYFYFFDEQAIRAFRLVK
SNKELWALCLYPMVSWLACTCLMQMKREKLTSTKTTTLLCLHQAALQALPLP
QLRDLCSLAAEGINQKTLFSPDDLRLKHGLDGAIIISTFLKMGILQHPILSYSFIHL

CFQEFPAAMSYVLEDEKGRKHSNCIIDLKLEAYIGHLFGASTTRFLGLLSDG
 EREMEINIFRSLQGNLMQWVPSLQLLQPHSLHCLYETRNKFLTOVMHFEE
 MCMVETIMELLAVCTFCIKFSRHVKLIQIIEGRQHSWSPMVLFRWVPVTDAYWO
 IILFVLKVRNLKELDLSGNSLSHSNAVKSILCKTLRPRCLLETBLRAGGLTAEDCKD
 LAFGLRANOTLLELDLUSFNVLTDGAKHLCQRLRPSCKLQRLIVSGELTSDCCQDL
 ASVLSASPLELDLQNNLDVGNRLHCEGRHPACKLIRGLDQTLTSDMRQELR
 ALQEKPOLLIIFSRKPSVMTTEGJDTGEMSNSTSLKRQLGSRERASHVAQANLK
 LLDVSKIFPIABIAESSPEVVPVELLCPSPASQDHLTKPLGTDDDDFWPGTGPVAT
 EVDVEKNLYRVHFPVAGSVRPNTGLCFVMEAVTVEIEFCVMDQFLGEINPOHSWM
 VAGPLLDIKAEPGAIEAVHLPHFVALQGGHVDTSLFQMAHFKEEGKLEKPARVLEHH
 IVLENFSPLGLVGLKMIHNAURFIPVTSVLLVIRHPEEVTFHYLIIPSDCSIRKE
 LELCYRSPEDQLFSEFYVGHLSGIRLQVKDKDBETLVMEALVPGDLMPPATLIPP
 ARIAVSPDPAQLLFVDFQREQLIARTVSEVLDKLHGLQVLSQEOERYVLAENTR
 PSQMRKLFSLSSQWDRKDKGLYOALKETHPHLIMELWEKSKGLLPLSS"

BASE COUNT 1011 a 1189 c 1209 g 881 t

ORIGIN

Query Match 78.8%; Score 4290; DB 9; Length 4290;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 4290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 523 ATGGCTGGCGGAGCTGGGGCGCGCTGGCTGTACTTGGAGTTCCTGAAGAGGAG 582
 DB 1 ATGGCTGGCGGAGCTGGGGCGCGCTGGCTGTACTTGGAGTTCCTGAAGAGGAG 60

QY 583 CTGAAGAGTTCAGCTTCTGTCTGCGCAATAAGCGCACTCCAGGAGCTCTTCGGGTGAG 642
 DB 61 CTGAAGAGTTCAGCTTCTGTCTGCGCAATAAGCGCACTCCAGGAGCTCTTCGGGTGAG 120

QY 643 ACACCGCTCAGCAGAGAGAGAGTGGAGTGGCGCTCGTACCTGGAGTGGGTCTGAG 702
 DB 121 ACACCGCTCAGCAGAGAGAGAGTGGAGTGGCGCTCGTACCTGGAGTGGGTCTGAG 180

QY 703 TATGGGAGGAGGCGGCTGGAGCTTAGCCCTCCATACCTGGGAGAGATGGGCTGAG 762
 DB 181 TATGGGAGGAGGCGGCTGGAGCTTAGCCCTCCATACCTGGGAGAGATGGGCTGAG 240

QY 763 TCACTGTGCGCCCAAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 822
 DB 241 TCACTGTGCGCCCAAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300

QY 823 AGTGAACCCACCTGGGGTCTCCAGCCAAACCACTCCACCGCAGTGTCTAATGCCCTG 882
 DB 301 AGTGAACCCACCTGGGGTCTCCAGCCAAACCACTCCACCGCAGTGTCTAATGCCCTG 360

QY 883 ATCCATGAATTCGCGCGGGGTGACCCAGGCTCAGAGAGAGGTTTGGAGACGCTG 942
 DB 361 ATCCATGAATTCGCGCGGGGTGACCCAGGCTCAGAGAGAGGTTTGGAGACGCTG 420

QY 943 CTTGACACATCTGGAGCGCGCTGGAGAGAAATCTCTGCTCTACCTCTACCAAGCTCTT 1002
 DB 421 CTTGACACATCTGGAGCGCGCTGGAGAGAAATCTCTGCTCTACCTCTACCAAGCTCTT 480

QY 1003 CCAAGCTCCCCAGACCATGAGTCTCCAAGCAGGAGTCAACCAAGCCGCCCATCCACA 1062
 DB 481 CCAAGCTCCCCAGACCATGAGTCTCCAAGCAGGAGTCAACCAAGCCGCCCATCCACA 540

QY 1063 GCAGTCTGGGAGCTGGGATCCCCACCTCAGCCAGCTAGCAGCCAGAGAGAGGAG 1122
 DB 541 GCAGTCTGGGAGCTGGGATCCCCACCTCAGCCAGCTAGCAGCCAGAGAGAGGAG 600

QY 1123 GCTCTGGGACCAATGGCTCTGGATGAACGTCAAGAAATTTACTACACAGAAATCAGA 1182
 DB 601 GCTCTGGGACCAATGGCTCTGGATGAACGTCAAGAAATTTACTACACAGAAATCAGA 660

QY 1183 GAAAGAGAGAGAGAGAAATCAGAGAAAGGAGGCGCCCATGGGAGCGGTGGTAGGAAAG 1242
 DB 661 GAAAGAGAGAGAGAGAAATCAGAGAAAGGAGGCGCCCATGGGAGCGGTGGTAGGAAAG 720

QY 1243 CCCCCAGCGGACACAGGCTCAGAGCCGACACACACACACACACACACACACACACAC 1302
 DB 721 CCCCCAGCGGACACAGGCTCAGAGCCGACACACACACACACACACACACACACACAC 780

QY 1303 GAGAGCCTCTGTTCACATGCGCCTGGAAAAATGAGGATTTTAAACCAAAAAATTCACAG 1362
 DB 781 GAGAGCCTCTGTTCACATGCGCCTGGAAAAATGAGGATTTTAAACCAAAAAATTCACAG 840

QY 1363 CTGCTACTTCTACAAAGACCTCACCCAGAGACCCAGATCCCTGGTCAAGAGAAGCTGG 1422
 DB 841 CTGCTACTTCTACAAAGACCTCACCCAGAGACCCAGATCCCTGGTCAAGAGAAGCTGG 900

QY 1423 CTTGATTAATGTGGAGGAAATCGAGGACATTTAAATTGAGATCAGAGACTTATTTGGCCCA 1482
 DB 901 CTTGATTAATGTGGAGGAAATCGAGGACATTTAAATTGAGATCAGAGACTTATTTGGCCCA 960

QY 1483 GSCCTGGATACCCAGAACCTTCGATAGTCACTATGTCAGGCGGCTGTCGTAATGGGAAG 1542
 DB 961 GSCCTGGATACCCAGAACCTTCGATAGTCACTATGTCAGGCGGCTGTCGTAATGGGAAG 1020

QY 1543 TCAACACCTGGCGCAGGAGTGAAGAGCCTGGGGAGAGCCAGCTGTATTTGGGACCGC 1602
 DB 1021 TCAACACCTGGCGCAGGAGTGAAGAGCCTGGGGAGAGCCAGCTGTATTTGGGACCGC 1080

QY 1603 TTCCAGCATGTCTTCTACTTTCAGCTGTCAGAGCTGGCCCAAGTCCCAAGGTGTGAGTCTC 1662
 DB 1081 TTCCAGCATGTCTTCTACTTTCAGCTGTCAGAGCTGGCCCAAGTCCCAAGGTGTGAGTCTC 1140

QY 1663 GCTGAGCTCATCGGAAAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCCTGTCT 1722
 DB 1141 GCTGAGCTCATCGGAAAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCCTGTCT 1200

QY 1723 AGCCAGAGCGGCTGCTTCTTCTCATCTCGATGGTGTAGATGAGCCAGGATGGTCTTGAG 1782
 DB 1201 AGCCAGAGCGGCTGCTTCTTCTCATCTCGATGGTGTAGATGAGCCAGGATGGTCTTGAG 1260

QY 1783 GAGCCGAGTTCGTGAGCTCTGTCTGACCTGGAGCCAGCCACAGCGCGGAGTGCATGCTG 1842
 DB 1261 GAGCCGAGTTCGTGAGCTCTGTCTGACCTGGAGCCAGCCACAGCGCGGAGTGCATGCTG 1320

QY 1843 GGCAGTTTGTGGGAAAAATATATCTCCGAGGATCTCTTCTGATCACGGCTCGGACC 1902
 DB 1321 GGCAGTTTGTGGGAAAAATATATCTCCGAGGATCTCTTCTGATCACGGCTCGGACC 1380

QY 1903 ACAGCTCTGCGAAGCTCATTTCTTTTGGAGAGGACAGTGTGGGTAGAGGTCCTGGGG 1962
 DB 1381 ACAGCTCTGCGAAGCTCATTTCTTTTGGAGAGGACAGTGTGGGTAGAGGTCCTGGGG 1440

QY 1963 TTCTCTGAGTCCAGAGAGGAAATATTTCTACAGATATTTTACAGATGAAAGGCAAGCA 2022
 DB 1441 TTCTCTGAGTCCAGAGAGGAAATATTTCTACAGATATTTTACAGATGAAAGGCAAGCA 1500

QY 2023 ATTAGAGCCTTTAGGTTGGTCAAAATCAAAAGAGCTCTGGGCCCTGTGTCTTGTGCC 2082
 DB 1501 ATTAGAGCCTTTAGGTTGGTCAAAATCAAAAGAGCTCTGGGCCCTGTGTCTTGTGCC 1560

QY 2083 TGGGTGTCTGGTGGCTGACCTTGTGCTGATGAGCAGATGAGCGGAGGAAAGAACTC 2142
 DB 1561 TGGGTGTCTGGTGGCTGACCTTGTGCTGATGAGCAGATGAGCGGAGGAAAGAACTC 1620

QY 2143 ACAGTGAATTCAGAGACCCACACACCTCTGTCTACATTTACCTTGGCCAGGCTCTCCAA 2202
 DB 1621 ACAGTGAATTCAGAGACCCACACACCTCTGTCTACATTTACCTTGGCCAGGCTCTCCAA 1680

QY 2203 GCTCAGCCTTTGGGACCCAGCTCAGAGACCTCTGTCTCTGGTGTCTGAGGAGATCTGG 2262
 DB 1681 GCTCAGCCTTTGGGACCCAGCTCAGAGACCTCTGTCTCTGGTGTCTGAGGAGATCTGG 1740

QY 2263 CAAAAAGAGACCTTTTTCAGTCCAGATGACCTCAGAGAGCATGGGTTAGATGGGCGCATC 2322
 DB 1741 CAAAAAGAGACCTTTTTCAGTCCAGATGACCTCAGAGAGCATGGGTTAGATGGGCGCATC 1800

QY 2323 ATCTCAGCTTCTTGAAGATGGGTATTTCTTCAAGAGCACCCCATCTCCCTCTGAGCTACAGC 2382
 DB 1801 ATCTCAGCTTCTTGAAGATGGGTATTTCTTCAAGAGCACCCCATCTCCCTCTGAGCTACAGC 1860

QY 2383 TTCATTACCTCTGTTCCTCAAGAGGTTTCTTTTGCAGCAATGTCTCTATGTCTTGGAGGATGAG 2442

Db	1861	TTCAATTCACCTCTGTGTTTCCAAAGAGATTCTTTTGCAGCAATGTCCTATGCTTTTGGAGGATGAG	1920
Qy	2443	AAGGGGAGAGGTAAACATTCTAAATTGCATCATAGATTGGAAAAGACGCTAGAAGCATAT	2502
Db	1921	AAGGGGAGAGGTAAACATTCTAATTGCATCATAGATTGGAAAAGACGCTAGAAGCATAT	1980
Qy	2503	GGAAATCATGGCTGTGTTTGGGGCATCAACACACAGTTTCTATTTGGGCCCTGTTTAAAGTAT	2562
Db	1981	GGAAATCATGGCTGTGTTTGGGGCATCAACACACAGTTTCTATTTGGGCCCTGTTTAAAGTAT	2040
Qy	2563	GAGGGGAGAGAGATGGAGAACATCTTTTCACTGCGCGGTGTCTCAGGGGAGGAACTTG	2622
Db	2041	GAGGGGAGAGAGATGGAGAACATCTTTTCACTGCGCGGTGTCTCAGGGGAGGAACTTG	2100
Qy	2623	ATGCAGTGGGTCCCGTCCCTGTCAGCTGCTGCTGCAGGCACACTCTCTGAGTGCCTCCAC	2682
Db	2101	ATGCAGTGGGTCCCGTCCCTGTCAGCTGCTGCTGCAGGCACACTCTCTGAGTGCCTCCAC	2160
Qy	2683	TGCTTGTA CGAGA CTGCGAA CAAAA CGTTT CCTGCA CACAAGTGATGGCCCCATTTT CGAAGAA	2742
Db	2161	TGCTTGTA CGAGA CTGCGAA CAAAA CGTTT CCTGCA CACAAGTGATGGCCCCATTTT CGAAGAA	2220
Qy	2743	ATGGGCATGTGTGTAGAAA CAGACATGGAGCTCTTATGTGTGCATTTTCTGCAATTAATTC	2802
Db	2221	ATGGGCATGTGTGTAGAAA CAGACATGGAGCTCTTATGTGTGCATTTTCTGCAATTAATTC	2280
Qy	2803	AGCGCCACGTGAAGAAGCTTTCAGCTGATTGAGGCGAGGCAGACAGATCAACATGAGGC	2862
Db	2281	AGCGCCACGTGAAGAAGCTTTCAGCTGATTGAGGCGAGGCAGACAGATCAACATGAGGC	2340
Qy	2863	CCCACCATGTTAGTCTCTGTTTCAGTGGGGTCCCAGTCA CAGATGCCTATTGGCAGATTCCTC	2922
Db	2341	CCCACCATGTTAGTCTCTGTTTCAGTGGGGTCCCAGTCA CAGATGCCTATTGGCAGATTCCTC	2400
Qy	2923	TTCTCCGTCTCAAGGTCA CCGAGAA CCTGAAAGAGCTGGA CTTAAGTGGAAC TCGCTG	2982
Db	2401	TTCTCCGTCTCAAGGTCA CCGAGAA CCTGAAAGAGCTGGA CTTAAGTGGAAC TCGCTG	2460
Qy	2983	AGCCACTCTGCAGTGAAGAGTCTTTGTAAGACCCTGAGAC CGCCCTGCCTCCCTCGAG	3042
Db	2461	AGCCACTCTGCAGTGAAGAGTCTTTGTAAGACCCTGAGAC CGCCCTGCCTCCCTCGAG	2520
Qy	3043	ACCTCGGGTTGGCTGGCTGTGGCCCTCACAGCTGAGGACTGCAAGGACCTTGGCTTTGGG	3102
Db	2521	ACCTCGGGTTGGCTGGCTGTGGCCCTCACAGCTGAGGACTGCAAGGACCTTGGCTTTGGG	2580
Qy	3103	CTGAGAGCCAA CAGAC CCGTAG CCGAGCTGGA CTTGAGCTTCAATGTGCTCA CGGATGCT	3162
Db	2581	CTGAGAGCCAA CAGAC CCGTAG CCGAGCTGGA CTTGAGCTTCAATGTGCTCA CGGATGCT	2640
Qy	3163	GGAGCCAAA CACTTTTGCAGACTGGA CAGCCGAGCTGCAAGCTACAGCGACTCGAG	3222
Db	2641	GGAGCCAAA CACTTTTGCAGACTGGA CAGCCGAGCTGCAAGCTACAGCGACTCGAG	2700
Qy	3223	CTGCTCAGCTGTGGCCCTCACGCTCTGA CTGCTCCAGGACCTGGGCTCTGTGCTTAGTGCC	3282
Db	2701	CTGCTCAGCTGTGGCCCTCACGCTCTGA CTGCTCCAGGACCTGGGCTCTGTGCTTAGTGCC	2760
Qy	3283	AGCCCCAGCCTGAAGAGGACTAGACCTGCA CGAGAACAA CTTGGATGACGTTGGCGTGCGA	3342
Db	2761	AGCCCCAGCCTGAAGAGGACTAGACCTGCA CGAGAACAA CTTGGATGACGTTGGCGTGCGA	2820
Qy	3343	CTGCTCTGTGAGGGGCTCAGGCATCTGCTGCGAAA CTCATACGCTTGGGGCTGGACAG	3402
Db	2821	CTGCTCTGTGAGGGGCTCAGGCATCTGCTGCGAAA CTCATACGCTTGGGGCTGGACAG	2880
Qy	3403	ACAACTCTGAGTGTATGATGAGCGCAGGAAC TGAAGGCGCTTGGAGCAGGAGAAACCTCAG	3462
Db	2881	ACAACTCTGAGTGTATGATGAGCGCAGGAAC TGAAGGCGCTTGGAGCAGGAGAAACCTCAG	2940
Qy	3463	CTGCTCATCTTTACGACACGGAAACCAAGTGTGATGACCCCTACTGAGGGGCTTGATACG	3522

2941	Db	CTGCTCAATCTTACGACGCGAAACCAAGTGTGATGACCCCTACTGAGGCGCTGGATACG	3000
3523	Qy	GGAGAGATGAGTAATAGCACATCCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGGCG	3582
3001	Db	GGAGAGATGAGTAATAGCACATCCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGGCG	3060
3583	Qy	GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCCTGGACGTGAGCAAGATCTTCCCAATT	3642
3061	Db	GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCCTGGACGTGAGCAAGATCTTCCCAATT	3120
3643	Qy	GCTGAGATTCCAGAGAAAGCTCCACAGAGTAGTACCGGTGGAACTCTTTGTGCGTGCCT	3702
3121	Db	GCTGAGATTCCAGAGAAAGCTCCACAGAGTAGTACCGGTGGAACTCTTTGTGCGTGCCT	3180
3703	Qy	TCTCCTGCCCTCTCAAGGGGACCTGCATACGAAGCCTTTGGGGACTCAGCATGACTTCTGG	3762
3181	Db	TCTCCTGCCCTCTCAAGGGGACCTGCATACGAAGCCTTTGGGGACTCAGCATGACTTCTGG	3240
3763	Qy	GGCCCAACGGGGCCTGTGGCTACTGAGGTAGTTGACAAAGAAAAGAACTTTGTACCGAGTT	3822
3241	Db	GGCCCAACGGGGCCTGTGGCTACTGAGGTAGTTGACAAAGAAAAGAACTTTGTACCGAGTT	3300
3823	Qy	CACCTCCCTGTAGCTGGCTCCTACCGCTGSCCAAACACGGGTCTCTGCTTTGTGATGAGA	3882
3301	Db	CACCTCCCTGTAGCTGGCTCCTACCGCTGSCCAAACACGGGTCTCTGCTTTGTGATGAGA	3360
3883	Qy	GAAGCGGTGACCGTTGAGATTGAATTTCTGTGTGTGGACCAAGTCTCTGGGTGAGATCAAC	3942
3361	Db	GAAGCGGTGACCGTTGAGATTGAATTTCTGTGTGTGGACCAAGTCTCTGGGTGAGATCAAC	3420
3943	Qy	CCACAGCACAGCTGGATGTGGCAGGSCCTCTGTCTGGACATCAAGGTGAGCCTGGAGCT	4002
3421	Db	CCACAGCACAGCTGGATGTGGCAGGSCCTCTGTCTGGACATCAAGGTGAGCCTGGAGCT	3480
4003	Qy	GTGGAAGCTGTGACCTCCCTCACTTTGTGGCTCTCCAAAGGGGCCATGTGGACACATCC	4062
3481	Db	GTGGAAGCTGTGACCTCCCTCACTTTGTGGCTCTCCAAAGGGGCCATGTGGACACATCC	3540
4063	Qy	CTGTTCCAAATGGCCCACTTTAAAGAGAGGGGATGCTCCTCGAAGAACCCAGCCAGGGTG	4122
3541	Db	CTGTTCCAAATGGCCCACTTTAAAGAGAGGGGATGCTCCTCGAAGAACCCAGCCAGGGTG	3600
4123	Qy	GAGCTGCATCACATAGTFTCTGGAACCCCAAGTCTTCCCTTTGGAGTCTCTCTGAAA	4182
3601	Db	GAGCTGCATCACATAGTFTCTGGAACCCCAAGTCTTCCCTTTGGAGTCTCTCTCTGAAA	3660
4183	Qy	ATGATCATTAATGCCCTGCGCTTCATTCGCCGTCACTCTGTGTGTGCTTTTACCACCGC	4242
3661	Db	ATGATCATTAATGCCCTGCGCTTCATTCGCCGTCACTCTGTGTGTGCTTTTACCACCGC	3720
4243	Qy	GTCCATCCTGAGGAAGTCACTTCCACCTCTACTGATCCCAAGTGAAGTCTGCTCATTCGG	4302
3721	Db	GTCCATCCTGAGGAAGTCACTTCCACCTCTACTGATCCCAAGTGAAGTCTGCTCATTCGG	3780
4303	Qy	AAGGAACCTGGAGCTCTGCTATCGAAGCCCTGGAGAACCAAGCTGTTTCTCGGAGTTCTAC	4362
3781	Db	AAGGAACCTGGAGCTCTGCTATCGAAGCCCTGGAGAACCAAGCTGTTTCTCGGAGTTCTAC	3840
4363	Qy	GTTGGCCACTTTGGGATCAGGGATCAGGCTGGAAGTGAAGACAAAGAGATGAGACTCTG	4422
3841	Db	GTTGGCCACTTTGGGATCAGGGATCAGGCTGGAAGTGAAGACAAAGAGATGAGACTCTG	3900
4423	Qy	GTGTGGGAGGCCTTGTGTGAACCAAGAGATCTCATGCTGCAACTCTCTGATCCCTCCA	4482
3901	Db	GTGTGGGAGGCCTTGTGTGAACCAAGAGATCTCATGCTGCAACTCTCTGATCCCTCCA	3960
4483	Qy	GCCCGCATAGCCGTACCTTCACTCTGGATGCCCGCAGTGTGCTGCACTTTGTGGACCAAG	4542
3961	Db	GCCCGCATAGCCGTACCTTCACTCTGGATGCCCGCAGTGTGCTGCACTTTGTGGACCAAG	4020
4543	Qy	TATCGAGACAGCTGATAGCCCGAGTGACATCGGTGAGGTTGCTTTGGACAAACTGCAT	4602
4021	Db	TATCGAGACAGCTGATAGCCCGAGTGACATCGGTGAGGTTGCTTTGGACAAACTGCAT	4080

QY 301 GTCTTGAGGCCCTCTGGGCTCCTCCCTCCCTCCCTGGCTTTTCTTCAACACTCCCTCTAT 360
Db 331 GTCTTGAGGCCCTCTGGGCTCCTCCCTCCCTCCCTGGCTTTTCTTCAACACTCCCTCTAT 390
QY 361 CGGCGTCTATCTGTAGGTGCTCCCTGGGATTTATAAAATCTGGGTTCCGAATGCTGAATAAGA 420
Db 391 CGGCGTCTATCTGTAGGTGCTCCCTGGGATTTATAAAATCTGGGTTCCGAATGCTGAATAAGA 450
QY 421 GACGCTAAGAGCAAGGCAAGGCAAGCAAGCACTGTCTCTGCTGCTCTGATACCTCACCAC 480
Db 451 GACGCTAAGAGCAAGGCAAGGCAAGCAAGCACTGTCTCTGCTGCTCTGATACCTCACCAC 510
QY 481 CTGGGAACATCCCCAGACACCTCTTAACCTCGGGACAGAGATGGCTGGCGAGCTGG 540
Db 511 CTGGGAACATCCCCAGACACCTCTTAACCTCGGGACAGAGATGGCTGGCGAGCTGG 570
QY 541 GGCCTGCTGGCTGTATCTTGGAGTTCTTGAAGAGGAGGAGCTGAAGAGTTTCAGCTT 600
Db 571 GGCCTGCTGGCTGTATCTTGGAGTTCTTGAAGAGGAGGAGCTGAAGAGTTTCAGCTT 630
QY 601 CTGCTCGCCAAATAAGCGCACTCCAGGAGCTCTTCGGGTGAGACACCCGCTCAGCCAGAG 660
Db 631 CTGCTCGCCAAATAAGCGCACTCCAGGAGCTCTTCGGGTGAGACACCCGCTCAGCCAGAG 690
QY 661 AAGACGAGTGGCATGGAGTGGCTCGTACTCTGGTGGCTCAGTATGGGGAGCAGCGGGCC 720
Db 691 AAGACGAGTGGCATGGAGTGGCTCGTACTCTGGTGGCTCAGTATGGGGAGCAGCGGGCC 750
QY 721 TGGGACCTAGCCCTCATACCTGGGAGCAGATGGGGCTGAGTCACTGTGCGCCCAAGCC 780
Db 751 TGGGACCTAGCCCTCATACCTGGGAGCAGATGGGGCTGAGTCACTGTGCGCCCAAGCC 810
QY 781 CAGGAGGGGAGGCGCACTCTCCCTCATTCCTTCCCTACAGCCCAAGTGAACCCACCTGGG 840
Db 811 CAGGAGGGGAGGCGCACTCTCCCTCATTCCTTCCCTACAGCCCAAGTGAACCCACCTGGG 870
QY 841 TCTCCAGCCAAACCCACCTCCACCGCAGTGCCTAAATGCCCTGGATCCATGAATTCGCGCG 900
Db 871 TCTCCAGCCAAACCCACCTCCACCGCAGTGCCTAAATGCCCTGGATCCATGAATTCGCGCG 930
QY 901 GGGTGACCCAGGGCTCAGAGAGAGGGTTTTGAGACAGCTGCCTGACACATCTTGGAGCC 960
Db 931 GGGTGACCCAGGGCTCAGAGAGAGGGTTTTGAGACAGCTGCCTGACACATCTTGGAGCC 990
QY 961 CGCTGGAGAGAAATCTCTGCTCAGTCTCTTACCAGCTCTTCCAGCTCCCGACACAT 1020
Db 991 CGCTGGAGAGAAATCTCTGCTCAGTCTCTTACCAGCTCTTCCAGCTCCCGACACAT 1050
QY 1021 GAGTCTCCAGCCAGGAGTCAACCAAGCCGCCACATCCACAGCAGTGTCTGGGAGCTGG 1080
Db 1051 GAGTCTCCAGCCAGGAGTCAACCAAGCCGCCACATCCACAGCAGTGTCTGGGAGCTGG 1110
QY 1081 GGATCCCCA CTTAGCCCGCTAGCA CCGCAGAGCAGAGGCTCTCTGGGACCCCAATGG 1140
Db 1111 GGATCCCCA CTTAGCCCGCTAGCA CCGCAGAGCAGAGGCTCTCTGGGACCCCAATGG 1170
QY 1141 CCTCTGGATGAACCTCAGGAATTTACTACAGAAATCAGAGAAAGAGAGAGAGAA 1200
Db 1171 CCTCTGGATGAACCTCAGGAATTTACTACAGAAATCAGAGAAAGAGAGAGAGAA 1230
QY 1201 TCAGAGAAAGGAGGCGCCCATCGGCGAGCGGTGGTAGGACGCCGCCACAGGGCGACACC 1260
Db 1231 TCAGAGAAAGGAGGCGCCCATCGGCGAGCGGTGGTAGGACGCCGCCACAGGGCGACACC 1290
QY 1261 AGCCTACGCCCCACACCAACCCATCGGAGCTTTCTGTGAGAGAGGCTCTGTTCACA 1320
Db 1291 AGCCTACGCCCCACACCAACCCATCGGAGCTTTCTGTGAGAGAGGCTCTGTTCACA 1350
QY 1321 TGGCCCTGGAAAAATGAGGATTTTAAACCAAAATTCACACAGCTGCTACTTACAAAGA 1380
Db 1351 TGGCCCTGGAAAAATGAGGATTTTAAACCAAAATTCACACAGCTGCTACTTACAAAGA 1410

QY 1381 CCTCACCCCAGAGCCAAAGATCCCTCTGTCAAGAGAAAGCTGGCCTGATATATGTGAGAG 1440
Db 1411 TCTCACCCCAGAGCCAAAGATCCCTCTGTCAAGAGAAAGCTGGCCTGATATATGTGAGAG 1470
QY 1441 AATCGAGACATTTAAATTGAGATCAGAGACTTATTTGGCCAGGCTCGATACCCCAAGAA 1500
Db 1471 AATCGAGACATTTAAATTGAGATCAGAGACTTATTTGGCCAGGCTCGATACCCCAAGAA 1530
QY 1501 CCTCGCATAGTCATCTCGAGGGGCTCTGGAATTGGAAATCAACACTGGCCAGGAG 1560
Db 1531 CCTCGCATAGTCATCTCGAGGGGCTCTGGAATTGGAAATCAACACTGGCCAGGAG 1590
QY 1561 GTGAAGAAAGCTGGGGAGAGGCGCAGCTGTATGGGAGCCGCTTCAGCAATGCTCTTAC 1620
Db 1591 GTGAAGAAAGCTGGGGAGAGGCGCAGCTGTATGGGAGCCGCTTCAGCAATGCTCTTAC 1650
QY 1621 TTCAGCTCAGAGAGCTGGCCAGTCCAAAGTGGTGAAGTCTCGCTGAGCTCATCGGAAA 1680
Db 1651 TTCAGCTCAGAGAGCTGGCCAGTCCAAAGTGGTGAAGTCTCGCTGAGCTCATCGGAAA 1710
QY 1681 GATGGGACAGCCACTCCCGCTCCCATTTAGACAGATCTCTTAGGCCAGAGCGGCTGCTC 1740
Db 1711 GATGGGACAGCCACTCCCGCTCCCATTTAGACAGATCTCTTAGGCCAGAGCGGCTGCTC 1770
QY 1741 TTCATCTCGATGGTGTAGATGAGCCAGGATGGGCTTTGAGGAGCCGAGTTCTGAGCTC 1800
Db 1771 TTCATCTCGATGGTGTAGATGAGCCAGGATGGGCTTTGAGGAGCCGAGTTCTGAGCTC 1830
QY 1801 TGTCTGCACTGGAGCAGCCACAGCCGGGAGTGCATCTGCTGGGAGTTGCTGGGAAA 1860
Db 1831 TGTCTGCACTGGAGCAGCCACAGCCGGGAGTGCATCTGCTGGGAGTTGCTGGGAAA 1890
QY 1861 ACTATACTTCCCGAGGCACTCTCTGATCAGCGCTCGGACCAAGCTCTGAGAACCTC 1920
Db 1891 ACTATACTTCCCGAGGCACTCTCTGATCAGCGCTCGGACCAAGCTCTGAGAACCTC 1950
QY 1921 ATTCCTCTTTGGAGCAGGCACTGTTGGTGAAGTCTCTGGGTTCTGAGTCCAGCAG 1980
Db 1951 ATTCCTCTTTGGAGCAGGCACTGTTGGTGAAGTCTCTGGGTTCTGAGTCCAGCAG 2010
QY 1981 AAGGAATATTTCTACAGATATTTACAGATGAAGGCAAGCAATTAGAGCTTTAGTTG 2040
Db 2011 AAGGAATATTTCTACAGATATTTACAGATGAAGGCAAGCAATTAGAGCTTTAGTTG 2070
QY 2041 GTCAAATCAAAACAGAGCTCTGGCCCTGTCTGCTTGTGCTGCTGCTGCTGCTGCTG 2100
Db 2071 GTCAAATCAAAACAGAGCTCTGGCCCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2130
QY 2101 TGCACTTGCTGATGAGCAGATGAAGGCAAGGAAACTCACTGACTTCCAAGACC 2160
Db 2131 TGCACTTGCTGATGAGCAGATGAAGGCAAGGAAACTCACTGACTTCCAAGACC 2190
QY 2161 ACCAACAACCTCTGTCTACATTTACCTTGGCCAGGCTCTCCAAAGCTCAGCCATGGAGCC 2220
Db 2191 ACCAACAACCTCTGTCTACTATTTACCTTGGCCAGGCTCTCCAAAGCTCAGCCATGGAGCC 2250
QY 2221 CAGCTCAGAGACTCTGCTCTCTGCTGCTGAGGCACTCTGCAAAAAGAACCTTTTC 2280
Db 2251 CAGCTCAGAGACTCTGCTCTCTGCTGCTGAGGCACTCTGCAAAAAGAACCTTTTC 2310
QY 2281 AGTCCAGATGACTCAGGAAGCATGGTTAGATGGGGCCATCATCTCCAACCTTTTGAAG 2340
Db 2311 AGTCCAGATGACTCAGGAAGCATGGTTAGATGGGGCCATCATCTCCAACCTTTTGAAG 2370
QY 2341 ATGGGTATTTCTTCAAGAGCAACCCATTCCTCTGAGCTACAGCTTATTCACCTCTGTTTC 2400
Db 2371 ATGGGTATTTCTTCAAGAGCAACCCATTCCTCTGAGCTACAGCTTATTCACCTCTGTTTC 2430
QY 2401 CAAGAGTTCTTTGACAGCAATGCTCTATGCTTTGAGGATGAGAGGGGAGAGGTAACAT 2460
Db 2431 CAAGAGTTCTTTGACAGCAATGCTCTATGCTTTGAGGATGAGAGGGGAGAGGTAACAT 2490
QY 2461 TCTAATTGTCATCATAGATTTTGGAAAAAGCGCTAGAAAGCATATGGAATACATGCGCTGTTT 2520

Db 2491 TCTAAATTGCATCATAGATTGGAAGAGACGCTAGAAGCATATGGAATAATATGCGCTGTCTT 2550
 QY 2521 GGGGCAATCAACACACAGTTTCTTATTTGGCGCTGTTAAGTGATGAGGGGAGAGAGATG 2580
 Db 2551 GGGGCAATCAACACACAGTTTCTTATTTGGCGCTGTTAAGTGATGAGGGGAGAGAGATG 2610
 QY 2581 GAGAAATCTTTCACTGCGCGCTGCTCAGGGGAGGAACTGATGCAAGTGGTCCCGTCC 2640
 Db 2611 GAGAAATCTTTCACTGCGCGCTGCTCAGGGGAGGAACTGATGCAAGTGGTCCCGTCC 2670
 QY 2641 CTGACGTGCTGTGAGGCAACACTCTCTGGAGTCCCTCCACTGCTTGTGACAGACTCGG 2700
 Db 2671 CTGACGTGCTGTGAGGCAACACTCTCTGGAGTCCCTCCACTGCTTGTGACAGACTCGG 2730
 QY 2701 AACAAAACTTCTGTACACAGTGATGGGCCATTTGGAAGAAATGGGATGTTGTAGAA 2760
 Db 2731 AACAAAACTTCTGTACACAGTGATGGGCCATTTGGAAGAAATGGGATGTTGTAGAA 2790
 QY 2761 ACAGACATGGAGCTCTTAGTGTGCACTTTCTGCAATTAATTCAGCCGCCACCTGGAAGAG 2820
 Db 2791 ACAGACATGGAGCTCTTAGTGTGCACTTTCTGCAATTAATTCAGCCGCCACCTGGAAGAG 2850
 QY 2821 CTTTACGTGATTTGAGGCGAGGCAAGCAAGATCAATGAGGCCCAACCATGATGTCCTG 2880
 Db 2851 CTTTACGTGATTTGAGGCGAGGCAAGCAAGATCAATGAGGCCCAACCATGATGTCCTG 2910
 QY 2881 TTTACGTGGGTCCAGCTCAGATGCTGCTATTTGGCAGATTTCTCTTCCGCTCCTCAAGTC 2940
 Db 2911 TTTACGTGGGTCCAGCTCAGATGCTGCTATTTGGCAGATTTCTCTTCCGCTCCTCAAGTC 2970
 QY 2941 ACCAGAAAACCTGAAGAGCTGACCTTAAGTGAAACTTCGCTGAGCCACTCTGCAGTGAAG 3000
 Db 2971 ACCAGAAAACCTGAAGAGCTGACCTTAAGTGAAACTTCGCTGAGCCACTCTGCAGTGAAG 3030
 QY 3001 AGTCTTTTGAAGACCTGTGAGACCCCTGCTGCTCTCTGAGACCCCTGCGGTTGCTGGC 3060
 Db 3031 AGTCTTTTGAAGACCTGTGAGACCCCTGCTGCTCTCTGAGACCCCTGCGGTTGCTGGC 3090
 QY 3061 TGTGGCTCAGACGTGAGACTGCAAGGACCTTGGCTTGGGCTGAGAGCCCAACGAGCC 3120
 Db 3091 TGTGGCTCAGACGTGAGACTGCAAGGACCTTGGCTTGGGCTGAGAGCCCAACGAGCC 3150
 QY 3121 CTGACCGAGCTGACCTGAGCTTCAATGTGCTCAGGATGCTGGAGCCAAAACACTTTGC 3180
 Db 3151 CTGACCGAGCTGACCTGAGCTTCAATGTGCTCAGGATGCTGGAGCCAAAACACTTTGC 3210
 QY 3181 CAGAGACTGAGA CAGCCGAGCTGCAAGCTACAGCGACTGCGAGCTGCTGAGCTGTGGCCTC 3240
 Db 3211 CAGAGACTGAGA CAGCCGAGCTGCAAGCTACAGCGACTGCGAGCTGCTGAGCTGTGGCCTC 3270
 QY 3241 AGCTCTGACTGCTGCCAGGACCTGGCTCTGTGCTTGTAGTCCAGCCCGAGCTGAAGGAG 3300
 Db 3271 AGCTCTGACTGCTGCCAGGACCTGGCTCTGTGCTTGTAGTCCAGCCCGAGCTGAAGGAG 3330
 QY 3301 CTAGACTGTCAGCAGAAACCTGAGTACGTTGGGCTGCGACTGCTGTGAGGGGCTC 3360
 Db 3331 CTAGACTGTCAGCAGAAACCTGAGTACGTTGGGCTGCGACTGCTGTGAGGGGCTC 3390
 QY 3361 AGGCAATCTGCTGCAAACTCATACGCTTGGGCTGGACAGACAACTCTGAGTGATGAG 3420
 Db 3391 AGGCAATCTGCTGCAAACTCATACGCTTGGGCTGGACAGACAACTCTGAGTGATGAG 3450
 QY 3421 ATGAGGCAAGAACTGAGGGCCCTGAGAGAGGAGAAACCTCAGTGTCTCATCTTCAGCAGA 3480
 Db 3451 ATGAGGCAAGAACTGAGGGCCCTGAGAGAGGAGAAACCTCAGTGTCTCATCTTCAGCAGA 3510
 QY 3481 CGGAAACCAAGTGTGATGATCCCTACTGAGGGCTGAGATACGGGAGAGATGATGATAGC 3540
 Db 3511 CGGAAACCAAGTGTGATGATCCCTACTGAGGGCTGAGATACGGGAGAGATGATGATAGC 3570
 QY 3541 ACATCCTCACTCAAGCGGAGAGACTCGGATCAGAGAGGCGGCTTCCCATGTGCTCAG 3600

3571 ACATCCTCACTCAAGCGGAGAGACTCGGATCAGAGAGGCGGCTTCCCATGTGCTCAG 3630
 QY 3601 GCTAATCTCAAACTCCTCGAGCTGAGCAAGATCTTCCAAATTGCTGAGATTGCA----- 3654
 Db 3631 GCTAATCTCAAACTCCTCGAGCTGAGCAAGATCTTCCAAATTGCTGAGATTGCGGCAAG 3690
 QY 3655 -----GAGGAAAGCTCCCGAGAGGTAGTACCGGTGGAATCTTGTGTGGCTTCTCTCT 3708
 Db 3691 AGCCACGAGGAAAGCTCCCGAGAGGTAGTACCGGTGGAATCTTGTGTGGCTTCTCTCT 3750
 QY 3709 GCCTCTCAAGGGAGCTGCATACGAAAGCTTTGGGACTGAGATGACTTCTTGGGGCCCC 3768
 Db 3751 GCCTCTCAAGGGAGCTGCATACGAAAGCTTTGGGACTGAGATGACTTCTTGGGGCCCC 3810
 QY 3769 ACGGGGCTGTGGCTGACTGAGGTAGTTCAGAAAGAAAGAACTTGTACCGAGTTCACTTC 3828
 Db 3811 ACGGGGCTGTGGCTGACTGAGGTAGTTCAGAAAGAAAGAACTTGTACCGAGTTCACTTC 3870
 QY 3829 CCTGTAGCTGGCTCCTACCGCTGGCCAAACACGGGTCTCTGCTTTGTGATGAGAGAAAG 3888
 Db 3871 CCTGTAGCTGGCTCCTACCGCTGGCCAAACACGGGTCTCTGCTTTGTGATGAGAGAAAG 3930
 QY 3889 GTGACCGTTGAGATTGAAATCTGTGTGTGGGACCAAGTTCTTGGGTGAGATCAA CCAAG 3948
 Db 3931 GTGACCGTTGAGATTGAAATCTGTGTGTGGGACCAAGTTCTTGGGTGAGATCAA CCAAG 3990
 QY 3949 CACAGCTGATGGTGGCAGGGCTCTGCTGGACATCAAGGCTGAGCCTGGAGCTGTGAA 4008
 Db 3991 CACAGCTGATGGTGGCAGGGCTCTGCTGGACATCAAGGCTGAGCCTGGAGCTGTGAA 4050
 QY 4009 GCTGTGACCTCCCTCACTTTGTGGCTCTCAAGGGGCTCATGTGGACACATCCCTGTTTC 4068
 Db 4051 GCTGTGACCTCCCTCACTTTGTGGCTCTCAAGGGGCTCATGTGGACACATCCCTGTTTC 4110
 QY 4069 CAAATGGCCCACTTTAAAGAGAGGGGATGCTCTGGAGAAAGCCAGCCAGGGTGGAGCTG 4128
 Db 4111 CAAATGGCCCACTTTAAAGAGAGGGGATGCTCTGGAGAAAGCCAGCCAGGGTGGAGCTG 4170
 QY 4129 CATCACATAGTCTTGGAAACCCAGAGCTTCTCCCTTGGGAGTCTCTCCTGAAATGATC 4188
 Db 4171 CATCACATAGTCTTGGAAACCCAGAGCTTCTCCCTTGGGAGTCTCTCCTGAAATGATC 4230
 QY 4189 CATATGCCCCCTGCGCTTCACTTCCCGTCACTCTGTGTGTGTTGCTTTTACCCGCGTCCAT 4248
 Db 4231 CATATGCCCCCTGCGCTTCACTTCCCGTCACTCTGTGTGTGTTGCTTTTACCCGCGTCCAT 4290
 QY 4249 CTTGAGGAAGTCACTTCCCACTCTA CTTGATCCCAAGTGAAGTCTGCTCCTTCCG----- 4302
 Db 4291 CTTGAGGAAGTCACTTCCCACTCTA CTTGATCCCAAGTGAAGTCTGCTCCTTCCGAAAGGCC 4350
 QY 4303 ----- 4302
 Db 4351 ATAGATGATCTAGAAATGAAATTCAGTTTGTGGAATTCACAGCCACCCCGCTGACC 4410
 QY 4303 ----- 4302
 Db 4411 CCACCTTATATGGGCTGCTTACACTGTGTGCTGGTCTGGTTGAGGATGCTGGAATA 4470
 QY 4303 -----AAGGAACTGGAGCTCTGTATCGAAGCCCTGAGAGAGACCAAGTCTTCTGGAG 4356
 Db 4471 CTCCCCAAGGAACTGGAGCTCTGTATCGAAGCCCTGAGAGAGACCAAGTCTTCTGGAG 4530
 QY 4357 TTTCTACGTTGGCCTTGGGATCAGGATCAGGCTGCAAGTGAAGACAAAGAAATGATCAG 4416
 Db 4531 TTTCTACGTTGGCCTTGGGATCAGGATCAGGCTGCAAGTGAAGTGAAGTGAAGTGAAG 4590
 QY 4417 ACTCTGTGTGGAGGCTTGGTGAAGCCAGGA 4449
 Db 4591 ACTCTGTGTGGAGGCTTGGTGAAGCCAGGA 4623

LOCUS AX089761 4422 bp DNA linear PAT 21-MAR-2001
DEFINITION Sequence 1 from Patent WO0116170.
ACCESSION AX089761
VERSION AX089761.1 GI:13443933
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Reed, J.C.
TITLE Card proteins involved in cell death regulation
JOURNAL Patent: WO 0116170-A 1 08-MAR-2001;
The Burnham Institute (US)
FEATURES
source
1..4422
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
CDS
1..4422
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC34899.1"
/db_xref="GI:13443934"
/translation="MAGGAWRLACYLFELKKELKBFOLLANKAHSSSSGETPAQ
PEKTSMEVASYLVIAQYGRQAWDLALHTWEQMLRSLCAQAEAGAGHSPSPSPSE
PHLGSPQSTAVLPMWTHLPAGCTQSGERRVLRPLDPTSGRWREISASLLYOAL
RSSPHQSPQSPNAPTAVLGSWGPSPOPSPAPROEAPGTOWPLDETSGIYVTE
IRERREKSEKGRPPAAVVGTPPQNTSLQPHHPWPSPRESLCSWPKNEPFO
KFTQLLLQPHRPSDDYVEENRHLIEIRDLFGPLDQTPRIIVLQ
AAGIGKSTLRAQKEAMGQLYGRFQHFYFSCRELAQSVLSLAELIGKDGATP
APIQTLSPERLLFLDGVDPGWVLQPPSPSELCHWSQPOPADALLSLGKLTLP
EASFLLTALVNLNLIPLSEQARWVEVLGFSESSRKEFYFVFTDERQAIAPRLVK
SNKELWALCLVPMWNLCTCLMQOMRKEKLTLSKTTTLCILHLVLAQALQAPLGP
QURDLSLAELGWKQTLFSPDDLKRLKGLDCAIISTFLMGILQHPPLPSYFIHL
COFEPPAANSVLEDEKGRKNSCIIIDLEKLEAYGHGLFGASTFLLGLLSDG
EREMENIFRCLSQSHNMQWPSLQQLQPHLSLHCLYETRNKFTLQVMAHFE
MGMCVETDMLLAVCTFCIKFSRHKVKKLIQIEGRQHRSTWSPVWLFRWVPVTDAYWQ
ILFSVLKVRNLKELDLSGNSLSHSAVSKLCTLRPRCLLETRLAGCLTAEDCKD
LAFGRANOTLFDLDSFNVLADAGAKHLQBLRPSCKLQBLQVSGLTSDCCODL
ASVLSASPSELKDLOQNNDDVGRLLCEGRHHPACKLIRGLQDOTTLUSDQELR
ALEQEKPLLIFSRKPSYMTPTGEDTGMENSTSLKRLGLSERAASHVAQNLK
LIDVSKIFPIAIAESSEPVVELLQVSPASQDLHTKPLGDDDDPWGPTGVAT
EVVDKEKNLYRVHFPVAGSYRWPNGLCFVMEAEVTELEFCVMDQFGEINPQHSWM
VAGPLLDIKAEPCAEVAHLPHFVALQGGHVDTSLFQMAHFKKEGMLLEKPARVLKH
IVLENPSPFLGWLKMHHNLRFPVTSVLLYHREVEVTEHLYLTPSDCSIRKA
IDLENKFPQVRIHKPPPLPTLMGCRYTVSGSGMSLILPKLELCITRSPGEDOLF
SEFYVHGLSGIRLQVKDKDETLWELVKPDLPATTLIPPARIAVPSPLDAPQL
LHFVDQYRQLIARVTSVVLKHLGOVLSQEQYERVLAEENTRPSQMRKLFSLQSW
DRCKDGLYQALKEPHLIMELWEKSGKGLLPLSS"
BASE COUNT 1043 a 1224 c 1240 g 915 t
ORIGIN
Query Match 76.2%; Score 4148; DB 6; Length 4422;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 4290; Conservative 0; Mismatches 0; Indels 132; Gaps 1;
QY 523 ATGGCTGGCGGAGCCTGGGCGCCCTGGCTGTACTTGGAGTTCCTGAAGAGGAGGAG 582
DB 1 ATGGCTGGCGGAGCCTGGGCGCCCTGGCTGTACTTGGAGTTCCTGAAGAGGAGGAG 60
QY 583 CTGAAGGAGTTCAGCTTCGTGCGCAATAAAGCGCACTCCAGGAGCTTCTGGGTGAG 642
DB 61 CTGAAGGAGTTCAGCTTCGTGCGCAATAAAGCGCACTCCAGGAGCTTCTGGGTGAG 120
QY 643 ACACCCGCTCAGCAGAGAGACGAGTGGCATGAGAGTGGCCCTCGTACCTGGTGGCTCAG 702
DB 121 ACACCCGCTCAGCAGAGAGACGAGTGGCATGAGAGTGGCCCTCGTACCTGGTGGCTCAG 180
QY 703 TATGGGAGCAGCGGCGCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGCGTGAGG 762
DB 181 TATGGGAGCAGCGGCGCTGGGACCTAGCCCTCCATACCTGGGAGCAGATGGGCGTGAGG 240

QY 763 TCACCTGTGGCCCAAGCCAGCCAGAGGGGAGGCGCACTCTCCCTCATTTCCCTTACAGCCCA 822
DB 241 TCACCTGTGGCCCAAGCCAGCCAGAGGGGAGGCGCACTCTCCCTCATTTCCCTTACAGCCCA 300
QY 823 AGTGAACCCCACTGGGTCTCCAGCCAA CCACCTCCACCGCAGTGTAAATGCGCTGG 882
DB 301 AGTGAACCCCACTGGGTCTCCAGCCAA CCACCTCCACCGCAGTGTAAATGCGCTGG 360
QY 883 ATCCATGAATTCGCCGGCGGGTGCACCCAGGCTCAGAGAGAGGGTTTGTGAGACAGCTG 942
DB 361 ATCCATGAATTCGCCGGCGGGTGCACCCAGGCTCAGAGAGAGGGTTTGTGAGACAGCTG 420
QY 943 CCTGACATCTGGACGCGCTGGAGAGAAATCTCTGCTCTACCTCTTACCAAGCTCTT 1002
DB 421 CCTGACATCTGGACGCGCTGGAGAGAAATCTCTGCTCTACCTCTTACCAAGCTCTT 480
QY 1003 CCAAGCTCCCAAGACCATGAGTCTCCAGCCAGGAGTCCACCAACGCCCCCACTCCACA 1062
DB 481 CCAAGCTCCCAAGACCATGAGTCTCCAGCCAGGAGTCCACCAACGCCCCCACTCCACA 540
QY 1063 GCAGTGTGGGAGCTGGGATCCCCACCTCAGCCAGCTAGCACCCAGAGAGGAGG 1122
DB 541 GCAGTGTGGGAGCTGGGATCCCCACCTCAGCCAGCTAGCACCCAGAGAGGAGG 600
QY 1123 GCTCTGGGACCAATGCTCTGATGAAACGTGAGAAATTTTATACAGAAATTCAGA 1182
DB 601 GCTCTGGGACCAATGCTCTGATGAAACGTGAGAAATTTTATACAGAAATTCAGA 660
QY 1183 GAAAGAGAGAGAGAAATCAGAGAAAGGCGGCCCCCATGGGCGCGTGTAGGAAAG 1242
DB 661 GAAAGAGAGAGAGAAATCAGAGAAAGGCGGCCCCCATGGGCGCGTGTAGGAAAG 720
QY 1243 CCCCCAGAGGCGCACACAGCCCTACAGGCCACACACCCATGGGAGCTTCTGTGAGA 1302
DB 721 CCCCCAGAGGCGCACACAGCCCTACAGGCCACACACCCATGGGAGCTTCTGTGAGA 780
QY 1303 GAGAGCTCTGTTTCCATGCGCCCTGGAAAAATAGGATTTTAA CCAAAAAATTCACACAG 1362
DB 781 GAGAGCTCTGTTTCCATGCGCCCTGGAAAAATAGGATTTTAA CCAAAAAATTCACACAG 840
QY 1363 CTGCTACTTCTACAAAGACCTACCCAGAGCCCAAGATCCCTGGTCAAGAGAGCTGG 1422
DB 841 CTGCTACTTCTACAAAGACCTACCCAGAGCCCAAGATCCCTGGTCAAGAGAGCTGG 900
QY 1423 CTGATTTATGTGGAGAGATCGAGGACATTTAA TTGAGATCAGAGACTTATTTGGCCCA 1482
DB 901 CTGATTTATGTGGAGAGATCGAGGACATTTAA TTGAGATCAGAGACTTATTTGGCCCA 960
QY 1483 GGCCTGGATACCAAGAACCTCGCATAGTCATATCTCAGGGGGCTGTGGAAATGGGAAG 1542
DB 961 GGCCTGGATACCAAGAACCTCGCATAGTCATATCTGCGAGGGGCTGTGGAAATGGGAAG 1020
QY 1543 TCAA CATCTGGCAGGAGGTGAAGAGCTCTGGGGAGAGGCGCAGCTGTATGGGAGCCGC 1602
DB 1021 TCAA CACTGGCAGGAGGTGAAGAGGCTCTGGGGAGAGGCGCAGCTGTATGGGAGCCGC 1080
QY 1603 TTCCAGACTGTCTTCTACTTCAGCTGCAGAGAGCTGGCCAGTCCAAAGTGTGAGTCTC 1662
DB 1081 TTCCAGACTGTCTTCTACTTCAGCTGCAGAGAGCTGGCCAGTCCAAAGTGTGAGTCTC 1140
QY 1663 GCTGAGCTCATCGGAAAAAGATGGGACAGCACTCCGGCTCCCATTTAGACAGATCCTGTCT 1722
DB 1141 GCTGAGCTCATCGGAAAAAGATGGGACAGCACTCCGGCTCCCATTTAGACAGATCCTGTCT 1200
QY 1723 AGGCCAGAGCGGTCTCTTCTATCTCGATGTGTAGATGAGCCAGATGGGTCTTGGAG 1782
DB 1201 AGGCCAGAGCGGTCTCTTCTATCTCGATGTGTAGATGAGCCAGATGGGTCTTGGAG 1260
QY 1783 GAGCCGAGTTCTGAGCTCTGTCTGACCTGGAGCCAGCCACAGCCGCGGATGCACCTG 1842
DB 1261 GAGCCGAGTTCTGAGCTCTGTCTGACCTGGAGCCAGCCAGCCGCGGATGCACCTG 1320
QY 1843 GGCAGTTTCTGGGAAAAAACTATATCTTCCGAGGATCTTCTCTGATCAGCGCTCGGACC 1902

Db	1321	GGCAGTTTGGGGGAAACTATACCTCCGAGGCATCTTCTGATCACGGCTGGACC	1380
Qy	1903	ACAGCTCTGAGAGAACTCACTTCCTTCTTTGGAGCAGCAGTGGGTAGAGTCTCTGGGG	1962
Db	1381	ACAGCTCTGAGAGAACTCACTTCCTTCTTTGGAGCAGCAGTGGGTAGAGTCTCTGGGG	1440
Qy	1963	TTCTCTGAGTCCAGCAGGAGGAATATTTCTACAGATATTTACAGATGAAGGCAAGCA	2022
Db	1441	TTCTCTGAGTCCAGCAGGAGGAATATTTCTACAGATATTTACAGATGAAGGCAAGCA	1500
Qy	2023	ATTAGAGCCTTTAGGTGGTCAATCAAAACAAAGAGCTCTGGGCCCTGTGTCTTTGGCCC	2082
Db	1501	ATTAGAGCCTTTAGGTGGTCAATCAAAACAAAGAGCTCTGGGCCCTGTGTCTTTGGCCC	1560
Qy	2083	TGGGTGCTGGTGGCTGCACTTGCCTGTATGACAGCAGATGAAGCGGAAGAAAACTC	2142
Db	1561	TGGGTGCTGGTGGCTGCACTTGCCTGTATGACAGCAGATGAAGCGGAAGAAAACTC	1620
Qy	2143	ACACTGACTTCCAAAGACCAACCAACCTCTGTCTACATTTACCTTGGCCAGGCTCTCCAA	2202
Db	1621	ACACTGACTTCCAAAGACCAACCAACCTCTGTCTACATTTACCTTGGCCAGGCTCTCCAA	1680
Qy	2203	GCTCAGCCATTTGGGACCCAGCTCAGAGACCTCTGCTCTCTGGCTGCTGAGGGCATCTGG	2262
Db	1681	GCTCAGCCATTTGGGACCCAGCTCAGAGACCTCTGCTCTCTGGCTGCTGAGGGCATCTGG	1740
Qy	2263	CAAAAAAGACCTTTTTCAGTCAGATGACCTCAGGAAGCATGGGTTAGATGGGGCCATC	2322
Db	1741	CAAAAAAGACCTTTTTCAGTCAGATGACCTCAGGAAGCATGGGTTAGATGGGGCCATC	1800
Qy	2323	ATCTCAGCTTCTTGAAGTGGTATCTTCAAGACACCCATCCCTCTGAGCTACAGC	2382
Db	1801	ATCTCAGCTTCTTGAAGTGGTATCTTCAAGACACCCATCCCTCTGAGCTACAGC	1860
Qy	2383	TTCAATTCACCTCTGTTTCCAAAGAGTTCTTTGAGCAATGCTCTATGCTTTGGAGGATGAG	2442
Db	1861	TTCAATTCACCTCTGTTTCCAAAGAGTTCTTTGAGCAATGCTCTATGCTTTGGAGGATGAG	1920
Qy	2443	AAGGGAGAGGTAAACATTTCTAATTCGATCATAGATTTGGAAGACGCTAGAGCATAT	2502
Db	1921	AAGGGAGAGGTAAACATTTCTAATTTGCATATAGATTTGGAAGACGCTAGAGCATAT	1980
Qy	2503	GGAAATACATGCTGTTTGGGGCATCAACACAGTTTCTTATGCGCTGTTAAAGTAT	2562
Db	1981	GGAAATACATGCTGTTTGGGGCATCAACACAGTTTCTTATGCGCTGTTAAAGTAT	2040
Qy	2563	GAGGGGAGAGAGATGGAGAACATCTTTCACTGCGGCTGTCTCAGGGGAGGAACCTG	2622
Db	2041	GAGGGGAGAGAGATGGAGAACATCTTTCACTGCGGCTGTCTCAGGGGAGGAACCTG	2100
Qy	2623	ATGCACTGGTCCCGTCCCTGACGCTGCTGTCAGCCCACTCTCTGAGTCCCTCCAC	2682
Db	2101	ATGCACTGGTCCCGTCCCTGACGCTGCTGTCAGCCCACTCTCTGAGTCCCTCCAC	2160
Qy	2683	TGCTTGTACGAGACTCGGAACAAACGTTCTTGACACAGTATGATGCCCCATTTGGAAGAA	2742
Db	2161	TGCTTGTACGAGACTCGGAACAAACGTTCTTGACACAGTATGATGCCCCATTTGGAAGAA	2220
Qy	2743	ATGGGCATGTGTAGAAAACAGATGGAGCTCTTATGTCAGCTTTCTGCAATTAATTC	2802
Db	2221	ATGGGCATGTGTAGAAAACAGATGGAGCTCTTATGTCAGCTTTCTGCAATTAATTC	2280
Qy	2803	AGCCGCCACGTGAAGAGCTTACGCTGATGAGGGCAGGACAGACAGATCAACATGGAGC	2862
Db	2281	AGCCGCCACGTGAAGAGCTTACGCTGATGAGGGCAGGACAGACAGATCAACATGGAGC	2340
Qy	2863	CCACCATGGTAGTCTGTTCAGGTGGGTCCAGTCCAGATGCCTATTGGCAGATTCTC	2922
Db	2341	CCACCATGGTAGTCTGTTCAGGTGGGTCCAGTCCAGATGCCTATTGGCAGATTCTC	2400
Qy	2923	TTCTCCGCTCTCAAGGTCAACGAAACCTGAAGGAGCTGGACCTAAGTGGAAATCTCGCTG	2982

Db	2401	TTCTCCGCTCTCAAGGTCAACAGAAACCTGAAGGAGCTGGACCTAAGTGGAAACTCGCTG	2460
Qy	2983	AGCACTCTGAGTGAAGAGTCTTTGTAAGACCCCTGACACGCCCTCGCTGCTCTCGTGG	3042
Db	2461	AGCACTCTGAGTGAAGAGTCTTTGTAAGACCCCTGACACGCCCTCGCTGCTCTCGTGG	2520
Qy	3043	ACCCTGCGGTTGGCTGGCTCAGCTGAGGACTGCAAGGACCTTTGCCTTTGGG	3102
Db	2521	ACCCTGCGGTTGGCTGGCTCAGCTGAGGACTGCAAGGACCTTTGCCTTTGGG	2580
Qy	3103	CTGAGAGCAACACGACCTGACCGAGCTGAGACCTGAGCTTCAATGTGCTCAGGATGCT	3162
Db	2581	CTGAGAGCAACACGACCTGACCGAGCTGAGACCTGAGCTTCAATGTGCTCAGGATGCT	2640
Qy	3163	GGAGCCAAACCTTTGCGAGAGCTGAGACGCGAGCTGCAAGCTACAGGACTGCGAG	3222
Db	2641	GGAGCCAAACCTTTGCGAGAGCTGAGACGCGAGCTGCAAGCTACAGGACTGCGAG	2700
Qy	3223	CTGGTCAGCTGGGCTCAGCTGCTGACCTGCTGCCAGGACCTGGCCCTCTGTAGTGCC	3282
Db	2701	CTGGTCAGCTGGGCTCAGCTGCTGACCTGCTGCCAGGACCTGGCCCTCTGTAGTGCC	2760
Qy	3283	AGCCCCAGCTGAGAGAGCTAGACCTGAGCAGAACTGCGATGAGCTGGGCTGGGA	3342
Db	2761	AGCCCCAGCTGAGAGAGCTAGACCTGAGCAGAACTGCGATGAGCTGGGCTGGGA	2820
Qy	3343	CTGCTCTGAGGGGCTCAGGCATCTGCTGCTGCAAACTCATACGCTGGGCTGGACACAG	3402
Db	2821	CTGCTCTGAGGGGCTCAGGCATCTGCTGCTGCAAACTCATACGCTGGGCTGGACACAG	2880
Qy	3403	ACAACTCTGAGTGATGAGAGCAGGAACTGAGGGCCCTGGAGCAGAGAACTCTCAG	3462
Db	2881	ACAACTCTGAGTGATGAGAGCAGGAACTGAGGGCCCTGGAGCAGAGAACTCTCAG	2940
Qy	3463	CTGCTCATCTTTCAGCAGACGAAACAAAGTGTGATGACCCCTACTAGGGCTGAGATAGC	3522
Db	2941	CTGCTCATCTTTCAGCAGACGAAACAAAGTGTGATGACCCCTACTAGGGCTGAGATAGC	3000
Qy	3523	GGAGAGTGTATAGACATCTCTCACTCAAGGGCAGAGACTCGGATCAGAGAGGGCG	3582
Db	3001	GGAGAGTGTATAGACATCTCTCACTCAAGGGCAGAGACTCGGATCAGAGAGGGCG	3060
Qy	3583	GCTTCCCATGTTGCTCAGGCTAACTCTCAAACTCTCTGAGCTGAGCAAGATCTTCCCAAT	3642
Db	3061	GCTTCCCATGTTGCTCAGGCTAACTCTCAAACTCTCTGAGCTGAGCAAGATCTTCCCAAT	3120
Qy	3643	GCTGAGATTGAGAGAAAGCTCCAGAGTAGTACCGGTGGAACTCTTGTGCTGCTGCT	3702
Db	3121	GCTGAGATTGAGAGAAAGCTCCAGAGTAGTACCGGTGGAACTCTTGTGCTGCTGCT	3180
Qy	3703	TCTCTGCTCTCAAGGGGACCTGCAATACGAGCTTTGGGGACTGAGGATGACTTCTGG	3762
Db	3181	TCTCTGCTCTCTCAAGGGGACCTGCAATACGAGCTTTGGGGACTGAGGATGACTTCTGG	3240
Qy	3763	GGCCCCACGGGCTGTGGCTACTGAGGTAGTTCACAAAGAAAGAACTTGTACCGAGTT	3822
Db	3241	GGCCCCACGGGCTGTGGCTACTGAGGTAGTTCACAAAGAAAGAACTTGTACCGAGTT	3300
Qy	3823	CACTTCCCTGTAGCTGGCTCTCAGCTGGCCCAACACGGGTCTCTGCTTTGTGATGAGA	3882
Db	3301	CACTTCCCTGTAGCTGGCTCTCAGCTGGCCCAACACGGGTCTCTGCTTTGTGATGAGA	3360
Qy	3883	GAAGGGTGACGGTTGAGATTGAATTTCTGTGTGGGACCAAGTTCTCTGGGTGAGATCAAC	3942
Db	3361	GAAGGGTGACGGTTGAGATTGAATTTCTGTGTGGGACCAAGTTCTCTGGGTGAGATCAAC	3420
Qy	3943	CCACAGCAGCTGATGGTGGCAGGGCTCTGCTGGACATCAAGGCTGAGCTGGAGCT	4002
Db	3421	CCACAGCAGCTGATGGTGGCAGGGCTCTGCTGGACATCAAGGCTGAGCTGGAGCT	3480
Qy	4003	GTGGAGCTGTGCACCTCCCTCCTTGTGGCTCTTCCAAAGGGGGCCATGTGAGACATCC	4062
Db	3481	GTGGAGCTGTGCACCTCCCTCCTTGTGGCTCTTCCAAAGGGGGCCATGTGAGACATCC	3540

Db 121 ACACCGCTCAGCCAGAGAGAAGACGAGTGGCATGGAGGTGGCTCTGCTACCTGTGGCTCAG 180
QY 703 TATGGGGAGCAGCGGCTTGGGACCTAGCCCTCATACCTGGGAGCAGATGGGCTGAGG 762
Db 181 TATGGGGAGCAGCGGCTTGGGACCTAGCCCTCATACCTGGGAGCAGATGGGCTGAGG 240
QY 763 TCACTGTGGCCCAAGCCCAAGGAGGGGAGGCCACTCTCCCTCATTTCCCTACAGCCCA 822
Db 241 TCACTGTGGCCCAAGCCCAAGGAGGGGAGGCCACTCTCCCTCATTTCCCTACAGCCCA 300
QY 823 AGTGAACCCCACTGGGGTCTCCAGCACAACCACTCCACCGCAGTGTAAATGCCCTGG 882
Db 301 AGTGAACCCCACTGGGGTCTCCAGCACAACCACTCCACCGCAGTGTAAATGCCCTGG 360
QY 883 ATCCATGAATTCGGCGGGGTGCACCCAGGGCTCAGAGAGAGGGTCTTGGAGACAGTG 942
Db 361 ATCCATGAATTCGGCGGGGTGCACCCAGGGCTCAGAGAGAGGGTCTTGGAGACAGTG 420
QY 943 CTGACACATCTGGACCGCGCTGGAGAGAAATCTCTGCCCTCACTCTCTACCAAGCTCTT 1002
Db 421 CTGACACATCTGGACCGCGCTGGAGAGAAATCTCTGCCCTCACTCTCTACCAAGCTCTT 480
QY 1003 CCAAGTCTCCCAAGCAGTCTCAAGCAGAGTCAACCAAGCCGCCCAATCCACA 1062
Db 481 CCAAGTCTCCCAAGCAGTCTCAAGCAGAGTCAACCAAGCCGCCCAATCCACA 540
QY 1063 GCAGTCTGGGAGCTGGGATCCCACTCAGCCAGCCTAGCAGCAGAGCAGGAG 1122
Db 541 GCAGTCTGGGAGCTGGGATCCCACTCAGCCAGCCTAGCAGCAGAGCAGGAG 600
QY 1123 GCTCTGGGACCCCAATGGGCTCTGGATGAAACGTCTAGGAAATTTACTACACAGAAATCAGA 1182
Db 601 GCTCTGGGACCCCAATGGGCTCTGGATGAAACGTCTAGGAAATTTACTACACAGAAATCAGA 660
QY 1183 GAAAGAGAGAGAGAGAAATCAGAGAAAGGAGGCCCAATGGGACGGTGTGAGAAAG 1242
Db 661 GAAAGAGAGAGAGAGAAATCAGAGAAAGGAGGCCCAATGGGACGGTGTGAGAAAG 720
QY 1243 CCCCACAGGCGCACACGAGCTACAGCCCAACCCACCCATGGGAGCCTTCTGTGAGA 1302
Db 721 CCCCACAGGCGCACACGAGCTACAGCCCAACCCACCCATGGGAGCCTTCTGTGAGA 780
QY 1303 GAGAGCCTCTGTTTCCAATGGGCTCTGGAAATGAGGATTTTAAACCAAAATTTACACAG 1362
Db 781 GAGAGCCTCTGTTTCCAATGGGCTCTGGAAATGAGGATTTTAAACCAAAATTTACACAG 840
QY 1363 CTGCTACTTCTACAAAGACCTCACCAGAGAGCCCAAGATCCCCTGGTCAAGAGAGCTGG 1422
Db 841 CTGCTACTTCTACAAAGACCTCACCAGAGAGCCCAAGATCCCCTGGTCAAGAGAGCTGG 900
QY 1423 CTTGATATGTGGAGAGAAATCGAGACATTTAAATTTGAGATCAGAGACTTATTTGGCCCA 1482
Db 901 CTTGATATGTGGAGAGAAATCGAGACATTTAAATTTGAGATCAGAGACTTATTTGGCCCA 960
QY 1483 GGCCTGGATACCAAGAACTCGCATAGTCAATCTCAGGGGCTCTGGAATTTGGGAAG 1542
Db 961 GGCCTGGATACCAAGAACTCGCATAGTCAATCTCAGGGGCTCTGGAATTTGGGAAG 1020
QY 1543 TCAACACTGGCCAGGAGTGAAGAGCCTGGGGAGAGCCAGCTGTATGGGAGCCGC 1602
Db 1021 TCAACACTGGCCAGGAGTGAAGAGCCTGGGGAGAGCCAGCTGTATGGGAGCCGC 1080
QY 1603 TTCCAGCATGCTTCTACTTACGTCAGAGAGCTGGCCAGTCCAGGTGGTGTGAGTCTC 1662
Db 1081 TTCCAGCATGCTTCTACTTACGTCAGAGAGCTGGCCAGTCCAGGTGGTGTGAGTCTC 1140
QY 1663 GCTGAGCTCATCGGAAAGATGGGACAGCACTCCGGCTCCCATTTAGACAGATCCTGTCT 1722
Db 1141 GCTGAGCTCATCGGAAAGATGGGACAGCACTCCGGCTCCCATTTAGACAGATCCTGTCT 1200
QY 1723 AGGCCAGAGCGGCTCTTCTCATCTCGATGTGTAGATGAGCCAGGATGGGCTTGGAG 1782
Db 1201 AGGCCAGAGCGGCTCTTCTCATCTCGATGTGTAGATGAGCCAGGATGGGCTTGGAG 1260

QY 1783 GAGCCGAGTCTGAGCTCTGTCTGCACTGGAGCCAGCCAGCCGCGGATGCACTGCTG 1842
Db 1261 GAGCCGAGTCTGAGCTCTGTCTGCACTGGAGCCAGCCAGCCGCGGATGCACTGCTG 1320
QY 1843 GGCAGTTTGTGGGAAATATATATTTCCGAGGATCTTCTCTGATCACTGCTGGACC 1902
Db 1321 GGCAGTTTGTGGGAAATATATATTTCCGAGGATCTTCTCTGATCACTGCTGGACC 1380
QY 1903 ACAGCTCTGCAGAACTCATTTCTTTTGGAGAGGCACTGGGTAGAGTCTCTGGGG 1962
Db 1381 ACAGCTCTGCAGAACTCATTTCTTTTGGAGAGGCACTGGGTAGAGTCTCTGGGG 1440
QY 1963 TTCTCTGAGTCCAGCAGGAAAGAAATATTTCTACAGATATTTTACAGATGAAAGGCAAGCA 2022
Db 1441 TTCTCTGAGTCCAGCAGGAAAGAAATATTTCTACAGATATTTTACAGATGAAAGGCAAGCA 1500
QY 2023 ATTAGAGCTTTAGGTTGGTCAAAATCAAAAGAGCTCTGGGCCCTGTGTCTTGTGCC 2082
Db 1501 ATTAGAGCTTTAGGTTGGTCAAAATCAAAAGAGCTCTGGGCCCTGTGTCTTGTGCC 1560
QY 2083 TGGGTCTCTGGCTGGCTGCACCTTGCCTGATGACAGAGATCAAGCGGAGGAAAAA CTC 2142
Db 1561 TGGGTCTCTGGCTGGCTGCACCTTGCCTGATGACAGAGATCAAGCGGAGGAAAAA CTC 1620
QY 2143 ACACCTGACTTCCAAAGACCAACCAACCTCTGTCTACATTTACCTTGGCCAGGCTCTCCAA 2202
Db 1621 ACACCTGACTTCCAAAGACCAACCAACCTCTGTCTACATTTACCTTGGCCAGGCTCTCCAA 1680
QY 2203 GCTCAGCATTGGGACCCAGCTCAGAGACCTCTGTCTCTGGTGTCTGAGGGCATCTGG 2262
Db 1681 GCTCAGCATTGGGACCCAGCTCAGAGACCTCTGTCTCTGGTGTCTGAGGGCATCTGG 1740
QY 2263 CAAAAAAGACCTTTTTCAGTCCAGATGACCTCAGGAGAGCATGGTTAGATGGGGCCATC 2322
Db 1741 CAAAAAAGACCTTTTTCAGTCCAGATGACCTCAGGAGAGCATGGTTAGATGGGGCCATC 1800
QY 2323 ATCTCACCTCTTGAAGATGGGTATTTCTTCAAGAGCACCCTCCTCTGAGCTACAGC 2382
Db 1801 ATCTCACCTCTTGAAGATGGGTATTTCTTCAAGAGCACCCTCCTCTGAGCTACAGC 1860
QY 2383 TTCAATTCACCTCTGTTTCCAAAGAGTCTTTTGGAGCAATGTCTATGTCTTGGAGATGAG 2442
Db 1861 TTCAATTCACCTCTGTTTCCAAAGAGTCTTTTGGAGCAATGTCTATGTCTTGGAGATGAG 1920
QY 2443 AAGGGAGAGGTAAACATTTCTAAATTTGCATCATAGATTTGGAAAGACCTGAGAGCATAT 2502
Db 1921 AAGGGAGAGGTAAACATTTCTAAATTTGCATCATAGATTTGGAAAGACCTGAGAGCATAT 1980
QY 2503 GGAATACATGGCTCTTTGGGGCATCAACCAACCTTTTCTTATTTGGGCTGTAAAGTGTAT 2562
Db 1981 GGAATACATGGCTCTTTGGGGCATCAACCAACCTTTTCTTATTTGGGCTGTAAAGTGTAT 2040
QY 2563 GAGGGGAGAGAGATGGAGAAACATTTTCACTCCGGCTGTCTCAGGGAGGAAACCTG 2622
Db 2041 GAGGGGAGAGAGATGGAGAAACATTTTCACTCCGGCTGTCTCAGGGAGGAAACCTG 2100
QY 2623 ATGAGTGGGTCCCGTCCCTGAGCTGTCTGAGCCACACTCTCTGAGAGTCCCTCCAC 2682
Db 2101 ATGAGTGGGTCCCGTCCCTGAGCTGTCTGAGCCACACTCTCTGAGAGTCCCTCCAC 2160
QY 2683 TGCTTGTACGAGACTCGGAACAAACCTTCTTGACACAGTGTGAGCCCATTTCCAGAA 2742
Db 2161 TGCTTGTACGAGACTCGGAACAAACCTTCTTGACACAGTGTGAGCCCATTTCCAGAA 2220
QY 2743 ATGGGCATGTGTGTAGAAACAGACATGGAGCTCTTGTGTGCACTTTCTGCAATTAATTC 2802
Db 2221 ATGGGCATGTGTGTAGAAACAGACATGGAGCTCTTGTGTGCACTTTCTGCAATTAATTC 2280
QY 2803 AGCCGCCAGTGTGAAGAGCTTCACTGATTTGAGGCGAGCAGCAGATCAACATGGAGC 2862
Db 2281 AGCCGCCAGTGTGAAGAGCTTCACTGATTTGAGGCGAGCAGCAGATCAACATGGAGC 2340

2863	Qy	CCACACATGCTAGTCTCTGTTTCAGGTGGGTCCCAAGTCAACAGATGCCATTATTTGGCAGATTCTC	2922
2341	Db	CCCACCATGCTAGTCTCTGTTTCAGGTGGGTCCCAAGTCAACAGATGCCATTATTTGGCAGATTCTC	2400
2923	Qy	TTCTCCGTCCTCAAGGTCACACAGAAACCTCAAGAGAGCTGCACCTAAAGTGAAACTCGCTG	2982
2401	Db	TTCTCCGTCCTCAAGGTCACACAGAAACCTCAAGAGAGCTGCACCTAAAGTGAAACTCGCTG	2460
2983	Qy	AGCCACTCTGCAGTGAAGAGTCTTTGTAAGACCCCTGAGACGCCCTCGCTGCCCTCTCTGGAG	3042
2461	Db	AGCCACTCTGCAGTGAAGAGTCTTTGTAAGACCCCTGAGACGCCCTCGCTGCCCTCTCTGGAG	2520
3043	Qy	ACCCTGGGTTTGGCTGTGGCTCACAAGCTGAGGACTGCAAGGACCTTGCTGCTTTGGG	3102
2521	Db	ACCCTGGGTTTGGCTGTGGCTCACAAGCTGAGGACTGCAAGGACCTTGCTGCTTTGGG	2580
3103	Qy	CTGAGAGCCAAACAGACCCCTGACCGAGCTCGAACCTGAGCTTCAATGTGCTCAACGGATGCT	3162
2581	Db	CTGAGAGCCAAACAGACCCCTGACCGAGCTCGAACCTGAGCTTCAATGTGCTCAACGGATGCT	2640
3163	Qy	GGAGCCAAAACACCTTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGCTTACAGCAGCTGCAG	3222
2641	Db	GGAGCCAAAACACCTTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGCTTACAGCAGCTGCAG	2700
3223	Qy	CTGGTGACTGTGGCTCAGTCTGACTGCTGCCAGGACCTGGCTCTGTGCTTAGTGCC	3282
2701	Db	CTGGTGACTGTGGCTCAGTCTGACTGCTGCCAGGACCTGGCTCTGTGCTTAGTGCC	2760
3283	Qy	AGCCCCAGCCTGAAGAGCTAGACCTGCACAGCAACCAACCTGATGACGTTTGGCTGCGA	3342
2761	Db	AGCCCCAGCCTGAAGAGCTAGACCTGCACAGCAACCAACCTGATGACGTTTGGCTGCGA	2820
3343	Qy	CTGCTCTGTGAGGGGCTCAGGCACTCTGCTCTGCAAACTCATAGCCCTGGGGCTGGACCG	3402
2821	Db	CTGCTCTGTGAGGGGCTCAGGCACTCTGCTCTGCAAACTCATAGCCCTGGGGCTGGACCG	2880
3403	Qy	ACAACCTGAGTGATGATGAGAGCAGGAACCTGAGGGCCCTGGAGCAGGAGAAACCTCAG	3462
2881	Db	ACAACCTGAGTGATGATGAGAGCAGGAACCTGAGGGCCCTGGAGCAGGAGAAACCTCAG	2940
3463	Qy	CTGCTCATCTTCAGCAGACCGAAACCAAGTGTGATGACCCCTACTGAGGGCCTGGATACG	3522
2941	Db	CTGCTCATCTTCAGCAGACCGAAACCAAGTGTGATGACCCCTACTGAGGGCCTGGATACG	3000
3523	Qy	GGAGAGATGAGTAAATAGCACATCCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGGCG	3582
3001	Db	GGAGAGATGAGTAAATAGCACATCCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGGCG	3060
3583	Qy	GCTTCCCATGTTGCTCAGGCTAAATCTCAAACTCCTGACGTGAGCAAGACTTCTCCCAATT	3642
3061	Db	GCTTCCCATGTTGCTCAGGCTAAATCTCAAACTCCTGACGTGAGCAAGACTTCTCCCAATT	3120
3643	Qy	GCTGAGATTGCAGAGGAAAGCTCCCGAGAGGTAGTACCGGTGGAACTCTTTGCGGTGCCCT	3702
3121	Db	GCTGAGATTGCAGAGGAAAGCTCCCGAGAGGTAGTACCGGTGGAACTCTTTGCGGTGCCCT	3180
3703	Qy	TCTCCTCCCTCTCAAGGGGACCTGCATACGAAGCTTTTGGGGACTGACGATGACTTCTCG	3762
3181	Db	TCTCCTCCCTCTCAAGGGGACCTGCATACGAAGCTTTTGGGGACTGACGATGACTTCTCG	3240
3763	Qy	GGCCCCACGGGCTGTGGCTACTGAGGTAGTTGACAAAAGAAAGAACTTGTACCGAGTT	3822
3241	Db	GGCCCCACGGGCTGTGGCTACTGAGGTAGTTGACAAAAGAAAGAACTTGTACCGAGTT	3300
3823	Qy	CACTTCCCTGTAGCTGSGCTCCTACCGCTGGCCCCAACACGGGTCTCTGCTTTGTGATGAGA	3882
3301	Db	CACTTCCCTGTAGCTGSGCTCCTACCGCTGGCCCCAACACGGGTCTCTGCTTTGTGATGAGA	3360
3883	Qy	GAAGCGGTGACCGTTTGAGATTGAAATCTGTGTGTGGGACCAAGTTCTCTGGGTGAGATCAAC	3942
3361	Db	GAAGCGGTGACCGTTTGAGATTGAAATCTGTGTGTGGGACCAAGTTCTCTGGGTGAGATCAAC	3420
3943	Qy	CCACAGCACAGCTGGATGGTGGCAGGGCCCTCTGCTGACATCAAGCGCTGAGCCTGGAGCT	4002

[illegible]

RESULT 10
AF229061
LOCUS
DEFINITION

AF229061	61	CTGAAGGAGTTC	CGCCAAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG	120
AF229061.1	GI:12656108			
Homosapiens (human)				
Homosapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
Chuz.-L., Pio.F., Xie,Z., Godsik,A. and Reed,J.C.				
NAC: an Apaf-1/Ced-4 family member regulates the cytochrome c pathway for apoptosis				
Unpublished				
2 (bases 1 to 4200)				
Chu,Z.-L. and Reed,J.C.				
Direct Submission				
Submitted (28-JAN-2000) Apoptosis and Signal Transduction, The Burnham Institute, 10901 North Torrey Pines Road, La Jolla, CA 92037, USA				
Location/Qualifiers				
1..4200				
/organism="Homo sapiens"				
/mol_type="mRNA"				
/db_xref="taxon:9606"				
/cell_type="activated Jurkat T cells"				
1..4200				
/gene="NAC"				
1..4200				
/gene="NAC"				
/notes="Apaf-1/Ced-4 family apoptosis regulator; contains NB-ARC domain, CARD domain, and leucine-rich repeats (LRR)"				
/codon_start=1				
/product="NAC-gamma splice variant"				
/protein_id="AAK00750.1"				
/db_xref="GI:12656109"				
/translation="WAGAGWRGLACYLEFLKKEELKEFQLLLANKAHSRSSGETPAQPKTSGMEVASYLVIAQYGEQRAWDALHTWEGMLRSLCAQEGAGHSPSPFVPSPEPHLSPSOSTAVLWPTTHELPAGCTGSGRRVLRLOPDTSGRRWREISASLLYQALRSPHSPSPQSNAPSTAVLGSWSPROPSPSLAPRQEARPTQWLDSTSGIYTEIRERREKSEKGRPWAAVVPDPQAHSTSLQPHHPWEPSPRESLSTWSPWQEDFNQKRTQLLKLQPHRSDQPLVKRSGQYVVEENRHLIEIRDLFGPLDQBPRIVLQGAAGIKSTLARQVKEAWGRQYLDGRFQHFVFSCELAQSKVVSIAELIKDGTATPAPIQTLSPERBLTILQGVDPGWVLQEPSELQSWSPQPADALLSLGLKTLIPASQELTARTALQNLIPSELEQARVEVLGFSSESRKEYFYRPTDQRAIRFLVKSNKELWALCVTLVQSLKCTLMQOMKEKELTSTKTTTTLCLHLAQLAQPLGRQRLDCLSLAEGIWOKKTLFSPDDLKRLGLDGLAISIFPLKMGILQEHPIPLSFYFHLCFQEFPAAMSYYLEDEKGRHNSCIDLEKTLBAYGLHGFAGSTFRFLGLLSDGGRFEMENIFCRLSQGRNLQMWVPSLQLLQPHLSLESLCHLYETRNKTLFTQVMAHPBEEGMCVTDMLNVLVCTFCIKFSPHVKKLQLEGRHRSPTSVMVLPRWVPTVDAYWQILFSLVTKMRELQDLSGSHSAVSKLCKTRPRCLLETURLAGCGLTAECDKDLAFGLRANGTLELDLSFNVLTDAGAKHLQRLQPSCKLQRLQVSCGLTSCCQDLASVLGASPKLELDLQNNLDDVGLRLCEGLRHPACKLIRLGRPSVMTPEGLDTGEMNSYSSLRKGLDSEARASHVAQANLKLDDVKIPTFAETAESSPEVVPVELLQVPSAPQDLTKPLGTDGTFWGPTEVVDKEKNLYRVHPFVAGSYRWPNTGLCFVMEAVTEIEFCVMPQDFLGEINPOHSMVAVGLLDIKAEPCGAEVHLPHFVALQGHVDTLSFQMAHPKEEGMLLEKPARVELHHI VLENPSFSPGLVLLKWIHNAIRLFPVTSQVLLLYHRVHPPEVTPHLYLIPSDCSIRKELELCYRSPGEDQLFSEFYVHLGSGIRLOVKOKDETLVWEALVAPGDMFATLIPPARLAVSPSDAPQLLHFVDQYREQILARVTSVEVVDKLHGVLSQRYRVERVLAENTRPSQWRKLFLSLQSWDRCKDGLYQALKETHPHILMELKEGSKGLPLSS"				
BASE COUNT	986 a	1167 c	1180 g	867 t
ORIGIN				
Query Match	75.3%	Score 4100;	DB 9;	Length 4200;
Best Local Similarity	97.9%;	Pred. No. 0;		
Matches 4200;	Conservative	0;	Mismatches	0;
			Indels	90;
			Gaps	1;
Qy	523	ATGGCTGGCGGAGCTGGGGCGCGCTTACTTGGAGTTCTTCTGAAGAGAGGAG	582	
Db	1	ATGGCTGGCGGAGCTGGGGCGCGCTTACTTGGAGTTCTTCTGAAGAGAGGAG	60	
Qy	583	CTGAAGGAGTTC	CGCCAAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG	642

1141	Db	GCTGAGCTCATCGGAAAGATGGGACAGCCACTCCGGCTCCCATTAGACAGATCCTGTCT	1200
1723	Qy	AGGCCAGAGCGGCTGCTCTTCATCTCGATGGTGTAGATGAGCCAGCATGGGTCTTGCAG	1782
1201	Db	AGGCCAGAGCGGCTGCTCTTCATCTCGATGGTGTAGATGAGCCAGCATGGGTCTTGCAG	1260
1783	Qy	GAGCCGAGTTCTGAGCTCTGTCTGCATCTGAGCCAGCCAGCCAGCCGCGGATGCACTGCTG	1842
1261	Db	GAGCCGAGTTCTGAGCTCTGTCTGCATCTGGAGCCAGCCAGCCGCGGATGCACTGCTG	1320
1843	Qy	GGCAGTTTGCTGGGGAACATATCTTCCTCCGAGGCATCCCTCTCTGATCAGGGCTCGGACC	1902
1321	Db	GGCAGTTTGCTGGGGAACATATCTTCCTCCGAGGCATCCCTCTCTGATCAGGGCTCGGACC	1380
1903	Qy	ACAGCTCTGCAGAACCTCATTTCTCTTTGGAGCAGGCAGTGTGGGTAGAGGTCTCTGGGG	1962
1381	Db	ACAGCTCTGCAGAACCTCATTTCTCTTTGGAGCAGGCAGTGTGGGTAGAGGTCTCTGGGG	1440
1963	Qy	TTCTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTCAAGATGAAGAGCGCAAGCA	2022
1441	Db	TTCTCTGAGTCCAGCAGGAAGGAATATTTCTTACAGATATTTCAAGATGAAGAGCGCAAGCA	1500
2023	Qy	ATTAGAGCTTTAGGTGGTCAAAATCAAAACAAAGAGCTCTGGGCCCTGTCTCTGTGSCCC	2082
1501	Db	ATTAGAGCTTTAGGTGGTCAAAATCAAAACAAAGAGCTCTGGGCCCTGTCTCTGTGSCCC	1560
2083	Qy	TGGGTGTCTGGCTGGGCTGCATTTGCTGTGATGCAGCAGATGAAGCGGAAGGAAAAATCT	2142
1561	Db	TGGGTGTCTGGCTGGGCTGCATTTGCTGTGATGCAGCAGATGAAGCGGAAGGAAAAATCT	1620
2143	Qy	ACACTGACTTCCAAAGACCAACAAACCTCTGTCTTCAATTACCTTGCCCAAGGCTCTCCAA	2202
1621	Db	ACACTGACTTCCAAAGACCAACAAACCTCTGTCTTCAATTACCTTGCCCAAGGCTCTCCAA	1680
2203	Qy	GCTCAGCCATTGGGACCCAGCTCAGAGACCTCTGCTCTCTGGGCTGTCTGAGGCGCATCTGG	2262
1681	Db	GCTCAGCCATTGGGACCCAGCTCAGAGACCTCTGCTCTCTGGGCTGTCTGAGGCGCATCTGG	1740
2263	Qy	CAAAAAAGAGCCCTTTTCAGTCCAGATGACCTCAGGAAAGATGGGTTAGATGGGGCCATC	2322
1741	Db	CAAAAAAGAGCCCTTTTCAGTCCAGATGACCTCAGGAAAGATGGGTTAGATGGGGCCATC	1800
2323	Qy	ATCTCCACCTTCTTGAAGATGGGTATTCTTCAAGAGCACCCCATCCCTCTGAGCTACAGC	2382
1801	Db	ATCTCCACCTTCTTGAAGATGGGTATTCTTCAAGAGCACCCCATCCCTCTGAGCTACAGC	1860
2383	Qy	TTCAATTACCTCTGTTTCCAAAGAGTTCTTTGACAGCAATGTCCTATGCTTGGAGGATGAG	2442
1861	Db	TTCAATTACCTCTGTTTCCAAAGAGTTCTTTGACAGCAATGTCCTATGCTTGGAGGATGAG	1920
2443	Qy	AAGGGAGAGGTAAACATTCTTAAATTGCAATCATAGATTGGAAAAAGACGCTAGAGCATAT	2502
1921	Db	AAGGGAGAGGTAAACATTCTTAAATTGCAATCATAGATTGGAAAAAGACGCTAGAGCATAT	1980
2503	Qy	GGAATACATAGCCCTGTTTGGGGCATCAACACAGCGTTTCTTATTGGGCTCTGTTAAGTGA	2562
1981	Db	GGAATACATAGCCCTGTTTGGGGCATCAACACAGCGTTTCTTATTGGGCTCTGTTAAGTGA	2040
2563	Qy	GAGGGGAGAGAGATGGAGAAATCTTTTCACTCCGCGCTGTCTCAGGGGAGGAACTG	2622
2041	Db	GAGGGGAGAGAGATGGAGAAATCTTTTCACTCCGCGCTGTCTCAGGGGAGGAACTG	2100
2623	Qy	ATGCAGTGGGTCCCGTCCCTGCAGCTGTCTGCAGCCACACTCTCTGAGTCCCTCCAC	2682
2101	Db	ATGCAGTGGGTCCCGTCCCTGCAGCTGTCTGCAGCCACACTCTCTGAGTCCCTCCAC	2160
2683	Qy	TGCTTGTACGAGACTCGGAAACAAACAGTTCTCTGACACAAAGTATGGGCCATTTTCEAAGAA	2742
2161	Db	TGCTTGTACGAGACTCGGAAACAAACAGTTCTCTGACACAAAGTATGGGCCATTTTCEAAGAA	2220
2743	Qy	ATGGGCATGTGTGTAGAAAACAGATGGAGCTTTAGTGTGCATCTTCTGCATTAATAATC	2802
2221	Db	ATGGGCATGTGTGTAGAAAACAGATGGAGCTTTAGTGTGCATCTTCTGCATTAATAATC	2280

Qy	2803	AGCGGCCACGTGAAGAGCTTTCAGCTGATTGAGGGCAGGCAGACAGATCAAAATGGAGC	2862
Db	2281	AGCGGCCACGTGAAGAGCTTTCAGCTGATTGAGGGCAGGCAGACAGATCAAAATGGAGC	2340
Qy	2863	CCACACATGCTAGTTCCTGTTTCAGGTGGTCCCAGTCCACAGATCCCTATTGGCAGATTCTC	2922
Db	2341	CCACACATGCTAGTTCCTGTTTCAGGTGGTCCCAGTCCACAGATCCCTATTGGCAGATTCTC	2400
Qy	2923	TTCTCCGTCTCAAGGTCAACAGAAACCTGAAGGAGCTGCACCTTAAGTGGAAATCTCGCTG	2982
Db	2401	TTCTCCGTCTCAAGGTCAACAGAAACCTGAAGGAGCTGCACCTTAAGTGGAAATCTCGCTG	2460
Qy	2983	AGCACCTCTGCAGTAAGAGTCTTTGTAAAGACCTTGAGACGCCCTCGCTGCCCTCTGGAG	3042
Db	2461	AGCACCTCTGCAGTAAGAGTCTTTGTAAAGACCTTGAGACGCCCTCGCTGCCCTCTGGAG	2520
Qy	3043	ACCTGGGTTGGCTGGCTGGCCCTCACAGCTGAGGACCTGCAAGGACCTTGCTGCTTTGGG	3102
Db	2521	ACCTGGGTTGGCTGGCTGGCCCTCACAGCTGAGGACCTGCAAGGACCTTGCTGCTTTGGG	2580
Qy	3103	CTGAGGCCAACACAGACCCCTGACCGAGCTGGACCTGAGCTTTCAATGTGCTCACGGATGCT	3162
Db	2581	CTGAGGCCAACACAGACCCCTGACCGAGCTGGACCTGAGCTTTCAATGTGCTCACGGATGCT	2640
Qy	3163	GGAGCCAAACACCTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCGACTGCAG	3222
Db	2641	GGAGCCAAACACCTTTGCCAGAGACTGAGACAGCCGAGCTGCAAGCTACAGCGACTGCAG	2700
Qy	3223	CTGGTCACTGTGGCCTCAAGTCTGACTGCTGCCAGGACCTGGCCTCTGTGCTTAGTGCC	3282
Db	2701	CTGGTCACTGTGGCCTCAAGTCTGACTGCTGCCAGGACCTGGCCTCTGTGCTTAGTGCC	2760
Qy	3283	AGCCCCAGCCTGAAGGAGCTTAGACCTGCACAGAAACAACTGGATGACGTTGGCGTCGGA	3342
Db	2761	AGCCCCAGCCTGAAGGAGCTTAGACCTGCACAGAAACAACTGGATGACGTTGGCGTCGGA	2820
Qy	3343	CTGCTCTGTGAGGGGCTCAGGCACTCCTGCTGCCAAACTCAATAGCCCTGGGGCTGGACCG	3402
Db	2821	CTGCTCTGTGAGGGGCTCAGGCACTCCTGCTGCCAAACTCAATAGCCCTGGGGCTGGACCG	2869
Qy	3403	ACAACTCTGATGATGAGATGAGCGAGGAACCTGAGGGCCCTGGAGCAGGAGAAACCTCAG	3462
Db	2870	-----	2869
Qy	3463	CTGCTCATCTTCACGACACGGAAACCAAGTGTATGATGACCCCTACTGAGGGCCTGGATACG	3522
Db	2870	-----GGAAACCAAGTGTATGATGACCCCTACTGAGGGCCTGGATACG	2910
Qy	3523	GGAGAGATGAGTAATAGCACATCTCTACTCAAGCGGCAGAGACTCGGATCAGAGAGGGCG	3582
Db	2911	GGAGAGATGAGTAATAGCACATCTCTACTCAAGCGGCAGAGACTCGGATCAGAGAGGGCG	2970
Qy	3583	GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCCTGGAAGTGGACAAAGATCTTCCCAATT	3642
Db	2971	GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCCTGGAAGTGGACAAAGATCTTCCCAATT	3030
Qy	3643	GCTGAGATTGACAGGAAAGCTCCACAGAGGTAGTACCGGTGGAACTCTTGTGGGTGCTC	3702
Db	3031	GCTGAGATTGACAGGAAAGCTCCACAGAGGTAGTACCGGTGGAACTCTTGTGGGTGCTC	3090
Qy	3703	TCTCTCCCTCTCAAGGGGACCTGCATACCAAGCTTTGGGGACTGACGATGACTTCTCG	3762
Db	3091	TCTCTCCCTCTCAAGGGGACCTGCATACCAAGCTTTGGGGACTGACGATGACTTCTCG	3150
Qy	3763	GGCCCCACGGGCCCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTT	3822
Db	3151	GGCCCCACGGGCCCTGTGGCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTT	3210
Qy	3823	CATTCCCTGTAGCTGGCTCTTACCGCTGGCCCAACACGGGTCTCTGCTTTGTGATGAGA	3882
Db	3211	CATTCCCTGTAGCTGGCTCTTACCGCTGGCCCAACACGGGTCTCTGCTTTGTGATGAGA	3270

QY	3883	GAAGCGGTGACCGTTGAGATTGAATTTCTGTGTGGACAGTTCCTGGGTGAGATCAAC	3942
Db	3271	GAAGCGGTGACCGTTGAGATTGAATTTCTGTGTGGACAGTTCCTGGGTGAGATCAAC	3330
QY	3943	CCACAGCAGCTGGATGGTGGCAGGGCTTCTGTGGACATCAAGCTGAGCTGGAGCT	4002
Db	3331	CCACAGCAGCTGGATGGTGGCAGGGCTTCTGTGGACATCAAGCTGAGCTGGAGCT	3390
QY	4003	GTGGAGCTGTGACCTCTCACTTTGTGGCTCTCCAAAGGGGCCATGTGGACATCC	4062
Db	3391	GTGGAGCTGTGACCTCTCACTTTGTGGCTCTCCAAAGGGGCCATGTGGACATCC	3450
QY	4063	CTGTTCCAAATGSCCACTTTAAAGAGGAGGGATGCTCTCTGAGAACGACCGAGGTTG	4122
Db	3451	CTGTTCCAAATGSCCACTTTAAAGAGGAGGGATGCTCTCTGAGAACGACCGAGGTTG	3510
QY	4123	GAGCTGCATCATAGTTCTTGAAACCCAGCTTCTCCCTTGGAGTCTCTCTGAA	4182
Db	3511	GAGCTGCATCATAGTTCTTGAAACCCAGCTTCTCCCTTGGAGTCTCTCTGAA	3570
QY	4183	ATGATCCATAATGSCCTGCGCTTCACTTCCCGTCACCTCTGTGTGTGTTTACCAACCG	4242
Db	3571	ATGATCCATAATGSCCTGCGCTTCACTTCCCGTCACCTCTGTGTGTGTTTACCAACCG	3630
QY	4243	GTCCATCTCTGAGGAAGTCACTTCCACCTCTACCTGATCCCAAGTGAAGTCTCTCAATCGG	4302
Db	3631	GTCCATCTCTGAGGAAGTCACTTCCACCTCTACCTGATCCCAAGTGAAGTCTCTCAATCGG	3690
QY	4303	AGGAAGCTGGAGCTCTGCTATCGAAGCCCTGAGAAAGACAGCTGTTCTGGAGTTCTAC	4362
Db	3691	AGGAAGCTGGAGCTCTGCTATCGAAGCCCTGAGAAAGACAGCTGTTCTGGAGTTCTAC	3750
QY	4363	GTGGCCACTTGGGATCAGGATCAGCTGCAAGTGAAGACAAAGATCAGACTCTG	4422
Db	3751	GTGGCCACTTGGGATCAGGATCAGCTGCAAGTGAAGACAAAGATCAGACTCTG	3810
QY	4423	GTGTGGAGGCGCTTGTGAAACACAGAGATCTCATCGCTGCAACTACTCTGATCCCTCA	4482
Db	3811	GTGTGGAGGCGCTTGTGAAACACAGAGATCTCATCGCTGCAACTACTCTGATCCCTCA	3870
QY	4483	GCCGCGATAGCGTACCTTCACTCTGGATGCCCCGACAGTGTGCTGAGCAGCAG	4542
Db	3871	GCCGCGATAGCGTACCTTCACTCTGGATGCCCCGACAGTGTGCTGAGCAGCAG	3930
QY	4543	TATCGAGAGCAGTATAGCCGAGTGACATCGGTGGAGTTGCTTGGACAACTGCAT	4602
Db	3931	TATCGAGAGCAGTATAGCCGAGTGACATCGGTGGAGTTGCTTGGACAACTGCAT	3990
QY	4603	GGACAGGTGCTGAGCCAGGAGCAGTACGAGAGGTGCTGGCTGAGAACACGAGGCCAGC	4662
Db	3991	GGACAGGTGCTGAGCCAGGAGCAGTACGAGAGGTGCTGGCTGAGAACACGAGGCCAGC	4050
QY	4663	CAGATCGGAAAGCTGTTTACGTTGAGCCAGTCTGGGACCGGAAGTGCAAAATGAGCTC	4722
Db	4051	CAGATCGGAAAGCTGTTTACGTTGAGCCAGTCTGGGACCGGAAGTGCAAAATGAGCTC	4110
QY	4723	TACCAAGCCCTGAAGAGAGCCCATCTCCTACCTCATTTATGGAATCTGGGAGAGGCGAG	4782
Db	4111	TACCAAGCCCTGAAGAGAGCCCATCTCCTACCTCATTTATGGAATCTGGGAGAGGCGAG	4170
QY	4783	AAAAAGGGACTCTCTGCACTCAGCAGCTGA	4812
Db	4171	AAAAAGGGACTCTCTGCACTCAGCAGCTGA	4200

RESULT 11
AX089763
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE

AX089763
Sequence 3 from Patent WO0116170.
AX089763
AX089763.1 GI:13443935
Homo sapiens (human)

4194 bp
DNA
linear
PAT 21-MAR-2001

ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	1 Reed, J.C.			
TITLE	Card proteins involved in cell death regulation			
JOURNAL	Patent: WO 0116170-A 3 08-MAR-2001; The Burnham Institute (US)			
FEATURES	Location/Qualifiers			
source	1. .4194			
	/organism="Homo sapiens"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:9606"			
CDS	1. .4194			
	/note="unnamed protein product"			
	/codon_start=1			
	/protein_id="CAC34900.1"			
	/db_xref="GI:13443936"			
	/translation="MAGAWRLACYLEFLKKEELKEPQLLLANKAHSRSSGETPPAQ PEKTSMEVSYLVAQGEQAWDLALHTMEQGLRSLCAQAQAGHSPFPYSPSE PHLGSPOPTSTAVLPMWHELPACTGSGERRVLRQLPDTSGRRWRISASLLYQAL PSSDHESPOESNAPTSTVLGWSGPPQPSLAPRQEAQPTQWPLDETSIGIYTE IRERERSEKGRPPMAAVGCTPQAHTSLQPHHPHPSVRESLCSLSTWPKNEDFNQ KFTQLLLQRPSPQDELVKRSPWVVEENRHLIEIRDLFGELDTQERIVILQG AAGIKSTLARQVKEAMRGQLYGDYFQHFYFSCRELAQSKVSLASLIGKDGATP APIRQLSRPERLLFILDVDEPGWVLPQEPSSELHWSQPPADALIGSLGKTLIP BASFLITARTALONLIPSLRQARWVLPSESSRKEFYFYRTDQRAIRAFRLVK SNKELMALVPMWYLACTCLMOQKREKELTLTSTTTTLLCLHLYLAQALQAOLPGP QLRDLCSLAESGIWQKTLFSPDDLKHGLDGLIISTFLKMGILQEHPIPLSYFIHL CFQFFAAMSYLEDEKRGKHSNCIIDLKTLBAYGIHGLFGASTTFLGLLSDEG EREMENPHCLSQGRNLQMWPSLQLLPHLSLESLCYETRNKLTLTQWMAHFE MGMCVETDEMLLVCTFCIFSRHVKKLQLEGRQHRSTWPTVLRFRWVPTDAYWQ ILFVLKVTNRLKELDLSNLSHSAVKSLCTLRPRCLLETLRRLACGLTAEDCKD LAFGRANQTLTDLDSFNLTDAGAKSLCQLRQPSCKLQRLQVSCGLTSDCCQDL ASVLASPSLEKLDLQNNLDDVGRLLCEGRHPACKLIIRKESWMTPTFGLDGTGM SNSTSLKRRQGLSERAASHVAQANLLKLDVSKIPTAIEAESSEPVVVELLGVPS PASQGDHLTKPLGDDDFGPTGPVATEVDKRNLYRVHPFVAGSVRWPNTGLCFVM REAVTEVEFCVMDQFLCEINPQHSWMVAGPLLDIKAPGVAEAVHLPHFVALQGHV DTSLFQWAHFKKEGMLEKPARVELHILLENPSFSLGVLKMHNLRIPTVSVV LLYHRVPEEVTPHLYLIPSDCSI RELELCYRSPGBDOLFSEFYVGHLSGIRLQVKD KKDETLVWEALVKFGDMLPATLIPPAKIAVPSPLDAPQLLHFFVDQYREQLIARVTS EVLWDLKHGQVLSQEQYERVLAENTRPSQMRKLFSLQSQSDWRCKDGLYQALKETHPH LIMELWEKSGKKGLPLSS"			
BASE COUNT	984 a	1167 c	1176 g	867 t
ORIGIN				
Query Match	74.9%; Score 4078; DB 6; Length 4194;			
Best Local Similarity	97.8%; Pred. No. 0;			
Matches 4194; Conservative	0; Mismatches 0; Indels 96; Gaps 2;			
QY	523	ATGGCTGGCGGAGCCTGGGGCGCCTGGCTGCTTACTTTGGAGTTCTCTGAAGAAGGAGGAG	582	
Db	1	ATGGCTGGCGGAGCCTGGGGCGCCTGGCTGCTTACTTTGGAGTTCTCTGAAGAAGGAGGAG	60	
QY	583	CTGAAGGAGTTCCAGCTTCTGCTCGCCAAATAAAGCGCACTCCAGGAGCTTTCGGGTGAG	642	
Db	61	CTGAAGGAGTTCCAGCTTCTGCTCGCCAAATAAAGCGCACTCCAGGAGCTTTCGGGTGAG	120	
QY	643	ACACCCGCTCAGCCAGAGAAAGACAGTGGCATGAGGTGGCTCTGTACTGTGGTCTCAG	702	
Db	121	ACACCCGCTCAGCCAGAGAAAGACAGTGGCATGAGGTGGCTCTGTACTGTGGTCTCAG	180	
QY	703	TATGGGAGCAGCGGGCCTGGGACCTAGCCCTCCATACCTGGGACAGATGGGGCTGAGG	762	
Db	181	TATGGGAGCAGCGGGCCTGGGACCTAGCCCTCCATACCTGGGACAGATGGGGCTGAGG	240	
QY	763	TCACCTGTGGCCCAAGCCAGGAGGGCGAGGCCACTCTCCCTCATTTCCCTTACAGCCCA	822	
Db	241	TCACCTGTGGCCCAAGCCAGGAGGGCGAGGCCACTCTCCCTCATTTCCCTTACAGCCCA	300	
QY	823	AGTGAACCCCACTGGGGTCTCCNAGCCCAACCCACCTCCACGGAGTGTCTAATGCCCTGG	882	
Db	301	AGTGAACCCCACTGGGGTCTCCNAGCCCAACCCACCTCCACGGAGTGTCTAATGCCCTGG	360	

QY 883 ATCCATGAATTGCCCGCGGGGTGCACCCAGGGCTCAGAGAGAGGGTTTGTAGACAGCTG 942
Db |||||
361 ATCCATGAATTGCCCGCGGGGTGCACCCAGGGCTCAGAGAGAGGGTTTGTAGACAGCTG 420
QY |||||
943 CTTGACACATCTGGAGCGCGCTGGAGAGAAATCTCTGCCCTCACTCTCTACCAAGCTCTT 1002
Db |||||
421 CTTGACACATCTGGAGCGCGCTGGAGAGAAATCTCTGCCCTCACTCTCTCTACCAAGCTCTT 480
QY |||||
1003 CCAAGCTCCCCAGACCATCAGTCTCCAAAGCCAGGAGTCAACCAAGCCGCCCCACATCCACA 1062
Db |||||
481 CCAAGCTCCCCAGACCATCAGTCTCCAAAGCCAGGAGTCAACCAAGCCGCCCCACATCCACA 540
QY |||||
1063 GCAGTCTGGGAGCTGGGAGTCCCACTCAGCCAGCTAGCACCCAGAGAGAGAGAG 1122
Db |||||
541 GCAGTCTGGGAGCTGGGAGTCCCACTCAGCCAGCTAGCACCCAGAGAGAGAGAG 600
QY |||||
1123 GCTCTGGGACCAATGGCTCTGGATGAACGTCAGGAATTTACTACACAGAAATCAGA 1182
Db |||||
601 GCTCTGGGACCAATGGCTCTGGATGAACGTCAGGAATTTACTACACAGAAATCAGA 660
QY |||||
1183 GAAAGAGAGAGAGAGAAATCAGAGAAAGCAGGCCCCCATGGCAGCGGTGGTAGGAAG 1242
Db |||||
661 GAAAGAGAGAGAGAGAAATCAGAGAAAGCAGGCCCCCATGGCAGCGGTGGTAGGAAG 720
QY |||||
1243 CCCCCACAGCGCACACCGCTTACAGCCCCCACACCCATGGGAGGCTTTCTGTGAGA 1302
Db |||||
721 CCCCCACAGCGCACACCGCTTACAGCCCCCACACCCATGGGAGGCTTTCTGTGAGA 780
QY |||||
1303 GAGAGCTCTGTTCCATGCGGCTCGGAAATAGGATTTTAAACCAAAATTCACACAG 1362
Db |||||
781 GAGAGCTCTGTTCCATGCGGCTCGGAAATAGGATTTTAAACCAAAATTCACACAG 840
QY |||||
1363 CTGCTACTTCTACAAAGACTCACCCAGAGCCAGAGTCCCTGTCAGAGAGAGCTGG 1422
Db |||||
841 CTGCTACTTCTACAAAGACTCACCCAGAGCCAGAGTCCCTGTCAGAGAGAGCTGG 900
QY |||||
1423 CTTGATTTATGTGGAGAGAAATCAGAGACATTTAAATTTAGATCAGAGACTTTATTTGGCCCA 1482
Db |||||
901 CTTGATTTATGTGGAGAGAAATCAGAGACATTTAAATTTAGATCAGAGACTTTATTTGGCCCA 960
QY |||||
1483 GGCCTGGATACCAAGAACTGTCATAGTCACTGACAGGGGCTGCTGGAAATTTGGGAAG 1542
Db |||||
961 GGCCTGGATACCAAGAACTGTCATAGTCACTGACAGGGGCTGCTGGAAATTTGGGAAG 1020
QY |||||
1543 TCAACACTGGCCAGGAGTGAAGAAAGCTGGGGAGAGGCCAGCTGATGGGACCCG 1602
Db |||||
1021 TCAACACTGGCCAGGAGTGAAGAAAGCTGGGGAGAGGCCAGCTGATGGGACCCG 1080
QY |||||
1603 TTCCAGCATGTCTTTACTTTACGTCAGAGAGCTGGCCCAAGTCCAAAGTGGTGAAGTCTC 1662
Db |||||
1081 TTCCAGCATGTCTTTACTTTACGTCAGAGAGCTGGCCCAAGTCCAAAGTGGTGAAGTCTC 1140
QY |||||
1663 GTGAGCTCATCGGAAAGATGGGACAGCCACTCCGGCTCCCAATTAGACAGATCTGTCT 1722
Db |||||
1141 GTGAGCTCATCGGAAAGATGGGACAGCCACTCCGGCTCCCAATTAGACAGATCTGTCT 1200
QY |||||
1723 AGCCAGAGGGCTGCTCTCATCTCGATGGTGTAGATGAGCCAGGATGGCTCTGAG 1782
Db |||||
1201 AGCCAGAGGGCTGCTCTCATCTCGATGGTGTAGATGAGCCAGGATGGGCTCTGAG 1260
QY |||||
1783 GAGCCAGGTTCTGAGTCTGTCTGCACTGGAGCCAGCCAGCCGCGGATGCACTGGCTG 1842
Db |||||
1261 GAGCCAGGTTCTGAGTCTGTCTGCACTGGAGCCAGCCAGCCGCGGATGCACTGGCTG 1320
QY |||||
1843 GGCAGTTTCTGGGAAAACTATCTTCCGAGGCACTCTTCTGATCAGGCTCGGACC 1902
Db |||||
1321 GGCAGTTTCTGGGAAAACTATCTTCCGAGGCACTCTTCTGATCAGGCTCGGACC 1380
QY |||||
1903 ACAGCTCTGAGAACCTCATCTCTTTGGAGCAGGCAAGTTGGGTAGAGTCTCTGGG 1962
Db |||||
1381 ACAGCTCTGAGAACCTCATCTCTTTGGAGCAGGCAAGTTGGGTAGAGTCTCTGGG 1440

QY 1963 TTCTCTGAGTCCAGCAGGAAAGAAATATTTCTACAGATATTTTCACAGATGAAAGGCAAGCA 2022
Db |||||
1441 TTCTCTGAGTCCAGCAGGAAAGAAATATTTCTACAGATATTTTCACAGATGAAAGGCAAGCA 1500
QY |||||
2023 ATTAGAGCCTTTAGTGGTCAAAATCAAAACAAAGAGCTCTGGGCCCTGTGTCTTGTGCC 2082
Db |||||
1501 ATTAGAGCCTTTAGTGGTCAAAATCAAAACAAAGAGCTCTGGGCCCTGTGTCTTGTGCC 1560
QY |||||
2083 TGGGTGTCTGGCTGGCCTGCACCTTTGCTGTGATGACAGCAGATGAAAGCGGAAAGAAAACCTC 2142
Db |||||
1561 TGGGTGTCTGGCTGGCCTGCACCTTTGCTGTGATGACAGCAGATGAAAGCGGAAAGAAAACCTC 1620
QY |||||
2143 ACATGATCTTCAAGACACACCAACCCCTCTGTCTACATTTACCTTGGCCAGGCTCTCCAA 2202
Db |||||
1621 ACATGATCTTCAAGACACACCAACCCCTCTGTCTACATTTACCTTGGCCAGGCTCTCCAA 1680
QY |||||
2203 GCTCAGCATTGGGACCCAGCTCAGAGACCTCTCTCTCTGGCTGCTGAGGGCATCTGG 2262
Db |||||
1681 GCTCAGCATTGGGACCCAGCTCAGAGACCTCTCTCTCTGGCTGCTGAGGGCATCTGG 1740
QY |||||
2263 CAAAAAAGAGACCTTTTTCAGTCCAGATGACCTCAGGAAAGCATGGTTTATAGTGGGCCATC 2322
Db |||||
1741 CAAAAAAGAGACCTTTTTCAGTCCAGATGACCTCAGGAAAGCATGGTTTATAGTGGGCCATC 1800
QY |||||
2323 ATCTCCACTTCTTGAAAGATGGGTATTCTTCAAGAGCACCCCATCTCTCTGAGCTACAGC 2382
Db |||||
1801 ATCTCCACTTCTTGAAAGATGGGTATTCTTCAAGAGCACCCCATCTCTCTGAGCTACAGC 1860
QY |||||
2383 TTCATTCACTCTGTTTCCAAAGAGTCTTTTGCAGCAATGTCCTATGCTCTTGAGAGATGAG 2442
Db |||||
1861 TTCATTCACTCTGTTTCCAAAGAGTCTTTTGCAGCAATGTCCTATGCTCTTGAGAGATGAG 1920
QY |||||
2443 AAGGGAGAGGTAACATTTCTAAATTTGCATCATAGATTTTGGAAAAAGACGCTAGAAGCATAT 2502
Db |||||
1921 AAGGGAGAGGTAACATTTCTAAATTTGCATCATAGATTTTGGAAAAAGACGCTAGAAGCATAT 1980
QY |||||
2503 GGAATACATGGCTCTTTGGGGCATCAAACACACTTTCTTATTTGGGCTGTAAAGTAT 2562
Db |||||
1981 GGAATACATGGCTCTTTGGGGCATCAAACACACTTTCTTATTTGGGCTGTAAAGTAT 2040
QY |||||
2563 GAGGGGAGAGAGATGGAGAACATCTTTCACTGCCGCTCTCTCAGGGGAGAACCTG 2622
Db |||||
2041 GAGGGGAGAGAGATGGAGAACATCTTTCACTGCCGCTCTCTCAGGGGAGAACCTG 2100
QY |||||
2623 ATGCACTGGCTCCCTCCCTGCACTGCTGCTGCAAGCCACACTCTCTGAGTCCCTCCAC 2682
Db |||||
2101 ATGCACTGGCTCCCTCCCTGCACTGCTGCTGCAAGCCACACTCTCTGAGTCCCTCCAC 2160
QY |||||
2683 TGCTTTGACGACTCGGAAACAAAACGTTCTCTGACACAAGTGAAGGCCCATTTTCAAGAA 2742
Db |||||
2161 TGCTTTGACGACTCGGAAACAAAACGTTCTCTGACACAAGTGAAGGCCCATTTTCAAGAA 2220
QY |||||
2743 ATGGGCATGTGTGTAGAAAACAGACATGAGCTCTTCTAGTGTGCACTTTCTGCAATTAATTC 2802
Db |||||
2221 ATGGGCATGTGTGTAGAAAACAGACATGAGCTCTTCTAGTGTGCACTTTCTGCAATTAATTC 2280
QY |||||
2803 AGCCGCCACGCTGAAGAGCTTCAGCTGATTGAGGCGAGCAGCAGATCAACATGGAGC 2862
Db |||||
2281 AGCCGCCACGCTGAAGAGCTTCAGCTGATTGAGGCGAGCAGCAGATCAACATGGAGC 2340
QY |||||
2863 CCCACCATGGTGTAGTCTCTTTCAGGTGGGTCCCAAGTCAAGATGSCCTATTGGCAGATTTCTC 2922
Db |||||
2341 CCCACCATGGTGTAGTCTCTTTCAGGTGGGTCCCAAGTCAAGATGSCCTATTGGCAGATTTCTC 2400
QY |||||
2923 TTCTCCGCTCTCAAGGTCACAGAAACCTGAAGGAGCTGGACCTAAGTGGAAACCTCGCTG 2982
Db |||||
2401 TTCTCCGCTCTCAAGGTCACAGAAACCTGAAGGAGCTGGACCTAAGTGGAAACCTCGCTG 2460
QY |||||
2983 AGCCACTCTGCAAGTCAAGTCTTTTGAAGCCCTGAGAGCCCTCGCTCGCTCTCTGGAG 3042
Db |||||
2461 AGCCACTCTGCAAGTCAAGTCTTTTGAAGCCCTGAGAGCCCTCGCTCGCTCTCTGGAG 2520
QY |||||
3043 ACCCTCGGCTTGGCTGGCTGTGGCTCTCAGCTGAGGACTGCAAGGACCTTGCCTTTGGG 3102

Db 2521 ||||| ACCCTCGGTTGGCTGGCTTGCCTCAGAGTGGAGCTGCAAGACCTTGTCTTTGGG 2580
Qy 3103 CTGAGAGCCAAACAGACCTTGACCGAGCTGGACCTGAGCTTCAATGTCTCAGGATGCT 3162
Db 2581 CTGAGAGCCAAACAGACCTTGACCGAGCTGGACCTGAGCTTCAATGTCTCAGGATGCT 2640
Qy 3163 GGAGCCAAACACCTTTGCCAGAGACTGAGACGCGGAGCTGCAAGCTACAGGGACTGCGAG 3222
Db 2641 GGAGCCAAACACCTTTGCCAGAGACTGAGACGCGGAGCTGCAAGCTACAGGGACTGCGAG 2700
Qy 3223 CTGGTCAAGTGTGGCCCTCAGCTCTGATCTGCTGCGCAGGACCTGGCTCTGTGTAGTGCC 3282
Db 2701 CTGGTCAAGTGTGGCCCTCAGCTCTGATCTGCTGCGCAGGACCTGGCTCTGTGTAGTGCC 2760
Qy 3283 AGCCCCAGCTGAGGAGCTGAGCTGCGAGCAAGCAACCTGGATGAGCTTGGCGTGCGA 3342
Db 2761 AGCCCCAGCTGAGGAGCTGAGCTGCGAGCAAGCAACCTGGATGAGCTTGGCGTGCGA 2820
Qy 3343 CTGCTCTGTGAGGGGCTCAGGCACTCTGCTGCAAACTCATACGCTGGGGCTGGACACAG 3402
Db 2821 CTGCTCTGTGAGGGGCTCAGGCACTCTGCTGCAAACTCATACGCT----- 2867
Qy 3403 ACAACTCTGATGATGAGATGAGGCAAGAACTGAGGGCCCTGGAGCAGGAAAACTCAG 3462
Db 2868 ----- 2867
Qy 3463 CTGCTCATCTTACAGCAGCGGAACCAAGTGTGATGACCCCTACTGAGGGCCCTGGATAGC 3522
Db 2868 -----GAAACCAAGTGTGATGACCCCTACTGAGGGCCCTGGATAGC 2907
Qy 3523 GGAGAGATGAGTAATAGCACTCCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGGCG 3582
Db 2908 GGAGAGATGAGTAATAGCACTCCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGGCG 2967
Qy 3583 GCTTCCCATGTTGCTCAGGCTAATCTCAAATCTCTGGAAGTGAAGATCTTCCCAATT 3642
Db 2968 GCTTCCCATGTTGCTCAGGCTAATCTCAAATCTCTGGAAGTGAAGATCTTCCCAATT 3027
Qy 3643 GCTGAGATTCAGAGAGAACTCCCGAGAGTAGTACCGGTGMACTCTGTGCGTGCT 3702
Db 3028 GCTGAGATTCAGAGAGAAAGCTCCCGAGAGTAGTACCGGTGMACTCTGTGCGTGCT 3087
Qy 3703 TCTCTGCTCTCAAGGGACCTGCACTCAAGAGCTTTGGGAGCTGACGATGACTTCTGG 3762
Db 3088 TCTCTGCTCTCAAGGGACCTGCACTCAAGAGCTTTGGGAGCTGACGATGACTTCTGG 3147
Qy 3763 GGCCCCACGGGCTGTGCTACTGAGGTAGTTGACAAAGAAAGAACTTTGACCGAGTT 3822
Db 3148 GGCCCCACGGGCTGTGCTACTGAGGTAGTTGACAAAGAAAGAACTTTGACCGAGTT 3207
Qy 3823 CACTTCCCTGTAGTGGCTCCTACCGCTGGCCCAACACGGGTCTCTGCTTTGTGATGAGA 3882
Db 3208 CACTTCCCTGTAGTGGCTCCTACCGCTGGCCCAACACGGGTCTCTGCTTTGTGATGAGA 3267
Qy 3883 GAAGCGGTGACCTTGCAGATTGAATCTGTGTGGACAGTTCCTGGGTGAGATCAAC 3942
Db 3268 GAAGCGGTGACCTTGCAGATTGAATCTGTGTGGACAGTTCCTGGGTGAGATCAAC 3327
Qy 3943 CCACAGCAGCTGGATGGTGGCAGGGCTCTGCTGGACATCAAGGCTGAGCCTGGAGCT 4002
Db 3328 CCACAGCAGCTGGATGGTGGCAGGGCTCTGCTGGACATCAAGGCTGAGCCTGGAGCT 3387
Qy 4003 GTGGAAGCTGTGCACCTCCCTCACTTTGTGGCTCTCCAAAGGGGCCATGTGACACATCC 4062
Db 3388 GTGGAAGCTGTGCACCTCCCTCACTTTGTGGCTCTCCAAAGGGGCCATGTGACACATCC 3447
Qy 4063 CTGTTCCAAATGGCCCACTTTAAAGAGGAGGGATGCTCTCTGAGAGCCAGCAGGGTG 4122
Db 3448 CTGTTCCAAATGGCCCACTTTAAAGAGGAGGGATGCTCTCTGAGAGCCAGCAGGGTG 3507
Qy 4123 GAGCTGCATCAGATGTTCTGGAACCCCGAGCTTCTCCCGCTGGAGTCTCTCTCTGAAA 4182
|||||

Db 3508 GAGCTGCATCATAGTCTTCTGGAACCCCGAGCTTCTCCCGCTGGGAGTCTCTCTGAAA 3567
Qy 4183 ATGATCCCAATATGCGCTCGCTTCAATCCCGCTCACTCTGTGGTGTCTCTTTACACCGC 4242
Db 3568 ATGATCCCAATATGCGCTCGCTTCAATCCCGCTCACTCTGTGGTGTCTTTACACCGC 3627
Qy 4243 GTCCATCTCTGAGGAGTCACTTCCACCTCTACCTGATCCCAAGTGAAGTCTCCATTGGG 4302
Db 3628 GTCCATCTCTGAGGAGTCACTTCCACCTCTACCTGATCCCAAGTGAAGTCTCCATTGG- 3686
Qy 4303 AAGGAACCTGAGGAGTCTGATCTGAGAGCCCTGGAGAGCCAGCTTCTCCGAGTCTTAC 4362
Db 3687 --GGAACCTGAGGAGTCTGATCTGAGAGCCCTGGAGAGCCAGCTTCTCCGAGTCTTAC 3744
Qy 4363 GTTGCCCACTTTGGGATCAGGATCAGGCTGCAAGTGAAGAGCAAGAAAGATGAGACTCTG 4422
Db 3745 GTTGCCCACTTTGGGATCAGGATCAGGCTGCAAGTGAAGAGCAAGAAAGATGAGACTCTG 3804
Qy 4423 GTTGGGAGGCTTTGGTGAACCCAGGAGATCTCATGCTGCAACTACTCTGATCCCTCA 4482
Db 3805 GTTGGGAGGCTTTGGTGAACCCAGGAGATCTCATGCTGCAACTACTCTGATCCCTCA 3864
Qy 4483 GCCGCATAGCGTACCTTCACTCTGATGATGATGATGATGATGATGATGATGATGATGATG 4542
Db 3865 GCCGCATAGCGTACCTTCACTCTGATGATGATGATGATGATGATGATGATGATGATGATG 3924
Qy 4543 TATCGAGAGCAGCTGATAGCCGAGTACATCGTGGAGGTTGTCTTGGACAAACTGCAT 4602
Db 3925 TATCGAGAGCAGCTGATAGCCGAGTACATCGTGGAGGTTGTCTTGGACAAACTGCAT 3984
Qy 4603 GGACAGTGTCTGAGCCAGGAGCAGTACGAGAGGTTGTGGTCTGAGAAACAGAGGCCAGC 4662
Db 3985 GGACAGTGTCTGAGCCAGGAGCAGTACGAGAGGTTGTGGTCTGAGAAACAGAGGCCAGC 4044
Qy 4663 CAGATCGGAAGCTGTTGAGCTTGGAGCCAGTCTGGGAGCCGAGTGAAGATGAGATC 4722
Db 4045 CAGATCGGAAGCTGTTGAGCTTGGAGCCAGTCTGGGAGCCGAGTGAAGATGAGATC 4104
Qy 4723 TACCAAGCCCTGAGGAGACCCATCCTCAGCTTATGGAACCTCTGGGAGAGGGCAGC 4782
Db 4105 TACCAAGCCCTGAGGAGACCCATCCTCAGCTTATGGAACCTCTGGGAGAGGGCAGC 4164
Qy 4783 AAAAGGAGCTCTGCGCACTCAGCAGCTGA 4812
Db 4165 AAAAGGAGCTCTGCGCACTCAGCAGCTGA 4194

RESULT 12
AF229062

LOCUS Homo sapiens NAC-delta splice variant (NAC) mRNA linear PRI 02-FEB-2001
DEFINITION Homo sapiens NAC-delta splice variant (NAC) mRNA, complete cds,
alternatively spliced.

ACCESSION AF229062
VERSION AF229062.1 GI:12656110

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 4332)
Chen, Z.-L., Pao, F., Xie, Z., Godzik, A. and Reed, J.C.

NAC, an Apaf-1/Ced-4 family member regulates the cytochrome c

pathway for apoptosis

Unpublished

2 (bases 1 to 4332)

Chen, Z.-L. and Reed, J.C.

Direct Submission

Submitted (28-JAN-2000) Apoptosis and Signal Transduction, The

Burnham Institute, 10901 North Torrey Pines Road, La Jolla, CA

92037, USA

Location/Qualifiers

1. 4332

/organism="Homo sapiens"

source

		/mol_type="mRNA" /db_xref="taxon:9606" /chromosome="11" /cell_type="activated Jurkat T cells" 1..4332 /gene="NAC" /gene="NAC" /note="Apaf-1/Ced-4 family apoptosis regulator; contains NB-ARC domain, CARD domain, and leucine-rich repeats (LRR)" /codon_start=1 /product="NAC-delta splice variant" /protein_id="AAK00751.1" /db_xref="GI:12656111" /translation="WAGGAWGRILACVLEFLKKEELKEFOLLILANKAHRSSSGETPAQ PEKTSMEVASVLVQYGEORAWDLALHTWEOMGLRSLCAOAEAGHSPSPVSPSE PHLGSPSPSTAVLMPWHELHPAGCTQGSERRVLRQLDPTSGRRWRHREISALLQYAL PSSPHESPSQSPNAPTSTAVLGSWSPSPQPSLAPRQEPAGTQWPLDETSGIYYTE IRERERESEKGRPPAAVVGTPPOAHTSLQPHHPWEPSPVRESLSTWPKNEDFNQ KFTOLLRLPHRPSQDPLVKRSWPDVYVEENRHLEIRDLFGPLDQEPRIIVLQG AAGIKGTLARQVKEAWRGOLYGRFOHVFYFSRELAQSKVSLAELIKGDKATP APIRILSRPERELLFLDGVDEPGWVLQEPSSSELCLHWSOPOPADALLGSLIGKTLIP EASFLITARTTALQNLILPSLEQARWVEVLGFSSESRKEYFYRYFDERQARAERLVK SNKELMALCVLPWWSLACTCLMOQMRKEKLTLSKTTTTLCLHLQAQLAOQLPGP QLRDLCSLAAGEIWKOTLFLSPDRLRKHLDGAIISTFLKMGILQEHPIPLSYFIHL CFQEPFAAMSYYLDEKGRKHSNCI IDLEKTLEAYGHGLFGASTTFLGLLSDGEG BREMENI PHCLRSOGRNLMOWPSLQLLLOPHLSLHCLYETRNKTLFTLVQMAHPEE MCMCVETDMELLVCTFCIKFSRHVKLQLLEGROHRSSTWSPTMVLFRWVPTDAYWQ ILF5VLKTRNLKELDSGLSHSAVSLCKTLRRPRCLLETILRLAGCLTAEDCKD LAFGLRANOTLTDELSDNVLTDAGAKHLCORLPQSPCKLQRLQLVSGLTSDCCDDL ASVLASPSLKELDLQNNLDVGRLLCEGLRHPACKLIRLQKPSVMTPTSEGLDTE MSNSTSLKROBLGSRERAA5HVAOANLKLVDVSKI FPIAEIAEBSSEVPVPEVLLCVF SPASQGLDHTKPLGTDDDDPWGPTGPVATEVDVKEKNLYRVHPVPAVSGSRWPNLTGCVF MREAVTVLEFCVNDQFLGEPINPOHSWVWVAGPLLDIKAEPGAVEAVHLPHFVALOGGH VDTSLFQMAHFEKEGMILLEKPARVELHVIHLENPSFSPUGVLKMLHNAIRLPIVTSY VLLYHRHPEEYTHLTPSPCSIRKADIDLEMKFQFVIRIHKPPLPLPLYMGCRITYV SSGSGMPLLEILPKELELCYRFGEDOLFSEFYVGHLSGIRLQVKKDETLVMEALY KPGDLMPATLLIPPARIAVPSPLDAPOLLHFDVQYREQLIARVTSVEVVDLKLHGQVL SOEOYERVLAEINTRPSQMRKFLPSLSQSWDRKCKDGLYQALYKETHPHLIMELWEKSGK GLPLSS"		BASE COUNT 1018 a 1202 c 1211 g 901 t ORIGIN	
gene	Db	361	ATCCATGAATTTCCCGCGGGGTGCACCAGGGCTCAGAGAAAGGTTTTTGGACAGCTG	420	
	Qy	943	CCTGACACATCTGGACGCGCTGGAGAGAAATCTCTGCCTCACTCTCTACCAAGCTCTT	1002	
CDS	Db	421	CCTGACACATCTGGACGCGCTGGAGAGAAATCTCTGCCTCACTCTCTACCAAGCTCTT	480	
	Qy	1003	COAAGCTCCCAGACCATGATGCTCTCAAGCCAGGAGTCAACCAAGCCGCCACATCCACA	1062	
	Db	481	CCAAGCTCCCAGACCATGATGCTCTCAAGCCAGGAGTCAACCAAGCCGCCACATCCACA	540	
	Qy	1063	GCAGTGTCTGGGAGCTGGGATCCCACCTCAGCCCAAGCCTAGCACCCAGAGAGCAGGAG	1122	
	Db	541	GCAGTGTCTGGGAGCTGGGATCCCACCTCAGCCCAAGCCTAGCACCCAGAGAGCAGGAG	600	
	Qy	1123	GCTCCTGGGACCCAAATGGCCTCTGGATGAAACGTCAGGAATTTACTACACAGAAATCAGA	1182	
	Db	601	GCTCCTGGGACCCAAATGGCCTCTGGATGAAACGTCAGGAATTTACTACACAGAAATCAGA	660	
	Qy	1183	GAAAGAGAGAGAGAAATTCAGAGAAAGCGAGGCCCCCATGGGAGCGGTGTGAGAACG	1242	
	Db	661	GAAAGAGAGAGAGAGAAATTCAGAGAAAGCGAGGCCCCCATGGGAGCGGTGTGAGAACG	720	
	Qy	1243	CCCCCAGAGGCGCACACAGCCTACAGCCCAACACCCATGGGAGCCTTCTGTGAGA	1302	
	Db	721	CCCCCAGAGGCGCACACAGCCTACAGCCCAACACCCATGGGAGCCTTCTGTGAGA	780	
	Qy	1303	GAGAGCCTCTGTTTCCACATGGCCCTGGGAAATAGAGATTTTAAACAAAATTTACACAG	1362	
	Db	781	GAGAGCCTCTGTTTCCACATGGCCCTGGGAAATAGAGATTTTAAACAAAATTTACACAG	840	
	Qy	1363	CTGCTACTTCTCAAGAAGCTCACCCCAAGAGCCCAAGATCCCCTGGTCAAGAGAGCTGG	1422	
	Db	841	CTGCTACTTCTCAAGAAGCTCACCCCAAGAGCCCAAGATCCCCTGGTCAAGAGAGCTGG	900	
	Qy	1423	CCTGATATGTGGAGGAGAAATCGAGGACATTTAATTTAGATCAGAGACTTATTTGGCCCA	1482	
	Db	901	CCTGATATGTGGAGGAGAAATCGAGGACATTTAATTTAGATCAGAGACTTATTTGGCCCA	960	
	Qy	1483	GSCCTGGATACCCAAAGAACCTCGCATAGTCTACTGTCAGGGGGCTGTGGAATTTGGGAG	1542	
	Db	961	GSCCTGGATACCCAAAGAACCTCGCATAGTCTACTGTCAGGGGGCTGTGGAATTTGGGAG	1020	
	Qy	1543	TCACACCTGGCCAGGAGGTGAAGGAGCCTGGGGAGAGGCGCAGCTGTATGGGACCGC	1602	
	Db	1021	TCACACCTGGCCAGGAGGTGAAGGAGCCTGGGGAGAGGCGCAGCTGTATGGGACCGC	1080	
	Qy	1603	TTCCAGCATGTCTTCTACTTACGTGCGAGAGCTGGGCCAGTCCAGGTGTGAGTCTC	1662	
	Db	1081	TTCCAGCATGTCTTCTACTTCTAGCTGCGAGAGCTGGGCCAGTCCAGGTGTGAGTCTC	1140	
	Qy	1663	GCTGAGCTCATCGGAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCCTGTCT	1722	
	Db	1141	GCTGAGCTCATCGGAAAGATGGGACAGCCACTCCGGCTCCCATTTAGACAGATCCTGTCT	1200	
	Qy	1723	AGGCCAGAGCGGCTCTCTTCTCATCTCGATGTGTAGATGAGCCAGGATGGGTCTTGAG	1782	
	Db	1201	AGGCCAGAGCGGCTCTTCTTCTCATCTCGATGTGTAGATGAGCCAGGATGGGTCTTGAG	1260	
	Qy	1783	GAGCCGAGTCTGAGCTGTGTGTGCACTGGAGCCAGCCACAGCCGCGGATGCACTGCTG	1842	
	Db	1261	GAGCCGAGTCTGAGCTGTGTGTGCACTGGAGCCAGCCACAGCCGCGGATGCACTGCTG	1320	
	Qy	1843	GGCAGTGTCTGGGAAACATATATCTCCGAGGATCCTTCTCTGATCA CGGCTCGGACC	1902	
	Db	1321	GGCAGTGTCTGGGAAACATATATCTCCGAGGATCCTTCTCTGATCA CGGCTCGGACC	1380	
	Qy	1903	ACAGCTCTGCAGAACTCATCTCTTCTTTGGAGCAGGCA CGTGTGGGTAGAGTCTCTGGG	1962	
	Db	1381	ACAGCTCTGCAGAACTCATCTCTTCTTTGGAGCAGGCA CGTGTGGGTAGAGTCTCTGGG	1440	
	Qy	1963	TTCTCTGAGTCCAGCAGGAAGAAATTTTCTACAGATATTTTCA CAGATGAAAGCGCAAGCA	2022	

Db 1441 TTCTCTGAGTCCAGCAGGAAGGAAATATTTCTACAGATATTTACAGATGAAAGGCAAGCA 1500
 QY 2023 ATTAGAGCCTTTAGGTGTGTCAAATCAAAACAAAGAGCTCTGGGCGCTGTGTGTCCTGCCC 2082
 Db 1501 ATTAGAGCCTTTAGGTGTGTCAAATCAAAACAAAGAGCTCTGGGCGCTGTGTGTCCTGCCC 1560
 QY 2083 TGGGTGTCTGCTGGTGGCTGCACTTCCCTGATGTCAGCAGATGAAGCGGAAGGAAACAACTC 2142
 Db 1561 TGGGTGTCTGCTGGTGGCTGCACTTCCCTGATGTCAGCAGATGAAGCGGAAGGAAACAACTC 1620
 QY 2143 ACACAGCTTCCAAAGACCAACCAACCTCTGTCTACATTAACCTTGGCCAGGCTCTCCAA 2202
 Db 1621 ACACAGCTTCCAAAGACCAACCAACCTCTGTCTACATTAACCTTGGCCAGGCTCTCCAA 1680
 QY 2203 GCTCAGCCATTTGGGACCCAGCTCAGAGACCTCTGTCTCTGGCTGCTGAGGGCATCTGG 2262
 Db 1681 GCTCAGCCATTTGGGACCCAGCTCAGAGACCTCTGTCTCTGGCTGCTGAGGGCATCTGG 1740
 QY 2263 CAAAAAAGACCCCTTTTCAGTCCAGATGACCTCAGGAAGCATGGGTAGATGGGGCCATC 2322
 Db 1741 CAAAAAAGACCCCTTTTCAGTCCAGATGACCTCAGGAAGCATGGGTAGATGGGGCCATC 1800
 QY 2323 ATCTCACCTTCTTGAAGATGGGTATTTCTCAAGAGCAACCCATCCTCTGAGCTACAGC 2382
 Db 1801 ATCTCACCTTCTTGAAGATGGGTATTTCTCAAGAGCAACCCATCCTCTGAGCTACAGC 1860
 QY 2383 TTCAATTCACCTCTGTGTTCCAAAGATTTCTTTCAGCAATGCTCTGATGCTTGGAGATGAG 2442
 Db 1861 TTCAATTCACCTCTGTGTTCCAAAGATTTCTTTCAGCAATGCTCTGATGCTTGGAGATGAG 1920
 QY 2443 AAGGGAGAGGTAAACATTTCTAATTCATCATAGATTTTGGAAAGACGCTAGAAGCATAT 2502
 Db 1921 AAGGGAGAGGTAAACATTTCTAATTCATCATAGATTTTGGAAAGACGCTAGAAGCATAT 1980
 QY 2503 GGAATACATGGCTGTGTTGGGATCAACACACAGCTTCTTATGCGCTGTGTAAGTAT 2562
 Db 1981 GGAATACATGGCTGTGTTGGGATCAACACACAGCTTCTTATGCGCTGTGTAAGTAT 2040
 QY 2563 GAGGGGAGAGAGATGAGAACATCTTTCTACTGCGGCTGCTCAGGGGAGGAACTG 2622
 Db 2041 GAGGGGAGAGAGATGAGAACATCTTTCTACTGCGGCTGCTCAGGGGAGGAACTG 2100
 QY 2623 ATGCACTGGCTCCGCTCCCTGACGCTGCTGTCAGCCACACTCTCTGAGTCCCTCCAC 2682
 Db 2101 ATGCACTGGCTCCGCTCCCTGACGCTGCTGTCAGCCACACTCTCTGAGTCCCTCCAC 2160
 QY 2683 TGTCTGTAAGAGACTCGGAACAAACCGTTCTTCTGACACAAAGTATGAGTCCCATTTCCGAAGAA 2742
 Db 2161 TGTCTGTAAGAGACTCGGAACAAACCGTTCTTCTGACACAAAGTATGAGTCCCATTTCCGAAGAA 2220
 QY 2743 ATGGGATCTGTGTAGAAACAGACATGGAGCTTTAGTGTGCACTTTCTGCAATTAATTC 2802
 Db 2221 ATGGGATCTGTGTAGAAACAGACATGGAGCTTTAGTGTGCACTTTCTGCAATTAATTC 2280
 QY 2803 AGCGGCACTGTGAAGAGCTTCACTGATTTAGGGGAGGAGCAGACAGATCAACATGGAGC 2862
 Db 2281 AGCGGCACTGTGAAGAGCTTCACTGATTTAGGGGAGGAGCAGACAGATCAACATGGAGC 2340
 QY 2863 CCCACCATGATGATCTGTTTCAAGTGGGTCCAGTCCAGATGATGCTTATTTGGCAGATTTCTC 2922
 Db 2341 CCCACCATGATGATCTGTTTCAAGTGGGTCCAGTCCAGATGATGCTTATTTGGCAGATTTCTC 2400
 QY 2923 TTCTCCGCTCTCAAGTTCACAGAAACCTTGAAGAGCTTGAACCTTGAAGAGCTTGAACCTTGAAG 2982
 Db 2401 TTCTCCGCTCTCAAGTTCACAGAAACCTTGAAGAGCTTGAACCTTGAAGAGCTTGAACCTTGAAG 2460
 QY 2983 AGCCACTCTGCAAGTGAAGAGCTTTTGAAGAGCTTGAAGAGCTTGAAGAGCTTGAAGAGCTTGAAG 3042
 Db 2461 AGCCACTCTGCAAGTGAAGAGCTTTTGAAGAGCTTGAAGAGCTTGAAGAGCTTGAAGAGCTTGAAG 2520
 QY 3043 ACCCTGCGGTGGCTGGCTGTGGCTTACAGCTGAGGACTGCAAGAGCTTGCCTTTGGG 3102
 Db 2521 ACCCTGCGGTGGCTGGCTGTGGCTTACAGCTGAGGACTGCAAGAGCTTGCCTTTGGG 2580

QY 3103 CTGAGAGCAACCAAGCCTTGACCGAGCTGACCTGAGCTTCAATGTGCTCAGGATGCT 3162
 Db 2581 CTGAGAGCAACCAAGCCTTGACCGAGCTGACCTGAGCTTCAATGTGCTCAGGATGCT 2640
 QY 3163 GGAGCAAAACCTTTGCCAGAGACTGAGACAGCGAGCTGCAAGCTACAGCACTGAG 3222
 Db 2641 GGAGCAAAACCTTTGCCAGAGACTGAGACAGCGAGCTGCAAGCTACAGCACTGAG 2700
 QY 3223 CTGGTCACTGTGGCTCAGCTGACTGCTGCGCAGGACCTGGCTCTGTGCTTGTAGTCC 3282
 Db 2701 CTGGTCACTGTGGCTCAGCTGACTGCTGCGCAGGACCTGGCTCTGTGCTTGTAGTCC 2760
 QY 3283 AGCCCAAGCTTGAAGAGCTAGACCTGACAGCAACCACTGGATGAGCTTGGCTGCGA 3342
 Db 2761 AGCCCAAGCTTGAAGAGCTAGACCTGACAGCAACCACTGGATGAGCTTGGCTGCGA 2820
 QY 3343 CTGCTCTGTGAGGGCTCAGGATCTCTGCTGCAAACTCATACGCTTGGGCTGAGACAG 3402
 Db 2821 CTGCTCTGTGAGGGCTCAGGATCTCTGCTGCAAACTCATACGCTGCTG----- 2869
 QY 3403 ACACTCTGAGTATGAGATGAGGAGGAACTGAGGGGCTTGGAGCAGGAGAAACCTCAG 3462
 Db 2870 ----- 2869
 QY 3463 CTGCTCATCTTTCAGCAGACGGAAACCAAGTGTGATGACCCCTTACTGAGGGCTGATACG 3522
 Db 2870 ----- 2910
 QY 3523 GGAGAGTATGATTAATGACATCTCTACTCAAGCGGAGAGACTCGGATCAGAGAGGCG 3582
 Db 2911 GGAGAGTATGATTAATGACATCTCTACTCAAGCGGAGAGACTCGGATCAGAGAGGCG 2970
 QY 3583 GCTTCCCATGTTGCTCAGGCTAACTCTAACTCTGAGAGTGAAGCAATCTTCCCAATT 3642
 Db 2971 GCTTCCCATGTTGCTCAGGCTAACTCTAACTCTGAGAGTGAAGCAATCTTCCCAATT 3030
 QY 3643 GCTGAGATTGAGAGGAAAGCTCCCGAGAGTGTACCGGTGGAACCTCTGTCGCTGCT 3702
 Db 3031 GCTGAGATTGAGAGGAAAGCTCCCGAGAGTGTACCGGTGGAACCTCTGTCGCTGCT 3090
 QY 3703 TCTCTGCTCTCAAGGGGACCTGCATACGAAGCTTTGGGAGCTGACGATGCTTCTG 3762
 Db 3091 TCTCTGCTCTCAAGGGGACCTGCATACGAAGCTTTGGGAGCTGACGATGCTTCTG 3150
 QY 3763 GGCCCAAGGGGCTGCTGCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTT 3822
 Db 3151 GGCCCAAGGGGCTGCTGCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTT 3210
 QY 3823 CACTTCCCTGTAGCTGGCTCCTACCGCTGGCCCAACCGGCTCTGCTGTTGTGATGAGA 3882
 Db 3211 CACTTCCCTGTAGCTGGCTCCTACCGCTGGCCCAACCGGCTCTGCTGTTGTGATGAGA 3270
 QY 3883 GAAAGCGTGACCGTTGAGATTGAACTTCTGTGTGGGACCAAGTTCCTGGGTGAGATCAAC 3942
 Db 3271 GAAAGCGTGACCGTTGAGATTGAACTTCTGTGTGGGACCAAGTTCCTGGGTGAGATCAAC 3330
 QY 3943 CCACAGCAGCTGATGATGGTGGCAGGCTCTGCTGAGACATCAAGGCTGAGGCTGGAGCT 4002
 Db 3331 CCACAGCAGCTGATGATGGTGGCAGGCTCTGCTGAGACATCAAGGCTGAGGCTGGAGCT 3390
 QY 4003 GTGGAAGCTGTGACCTCCCTCCTCTGCTGCTCCAAAGGGGCGCATGTGACACATCC 4062
 Db 3391 GTGGAAGCTGTGACCTCCCTCCTCTGCTGCTCCAAAGGGGCGCATGTGACACATCC 3450
 QY 4063 CTGTTCCAAATGGCCCACTTTAAAGAGAGGGGATGCTCTCTGGAGAAACCAAGGAGGTG 4122
 Db 3451 CTGTTCCAAATGGCCCACTTTAAAGAGAGGGGATGCTCTCTGGAGAAACCAAGGAGGTG 3510
 QY 4123 GAGCTGCAATCAGATGTTCTGGAACCCAGCTTCTCCCTTGGGAGTCTCTCTGAAA 4182
 Db 3511 GAGCTGCAATCAGATGTTCTGGAACCCAGCTTCTCCCTTGGGAGTCTCTCTGAAA 3570

Db 481 CCAAGCTCCCGACCATGATCTCCAGCCAGGAGTCAACCAAGCCGCCCATCCACA 540
 Qy 1063 GCAGTCTCGGGAGCTGGGATCCCACTCTAGCCAGCCTAGCACCACGAGCAGGAG 1122
 Db 541 GCAGTCTCGGGAGCTGGGATCCCACTCTAGCCAGCCTAGCACCACGAGCAGGAG 600
 Qy 1123 GCTCTCGGACCAATGGCTCTGGATGAAAGCTCAGGAATTTACTACAGAAATCAGA 1182
 Db 601 GCTCTGGACCAATGGCTCTGGATGAAAGCTCAGGAATTTACTACAGAAATCAGA 660
 Qy 1183 GAAAGAGAGAGAGAAATCAGAGAAAGGAGGCCCTCCATGGGACGCGTGTGGAACG 1242
 Db 661 GAAAGAGAGAGAGAAATCAGAGAAAGGAGGCCCTCCATGGGACGCGTGTGGAACG 720
 Qy 1243 CCCCCACGGCCACACAGCCTACAGCCCCCACCACCCATGGGAGCCTTCTGTGAGA 1302
 Db 721 CCCCCACGGCCACACAGCCTACAGCCCCCACCACCCATGGGAGCCTTCTGTGAGA 780
 Qy 1303 GAGAGCCTCTGTTCCACATGGCCCTGGAATAATGAGGATTTAAACCAAAAATTCACACAG 1362
 Db 781 GAGAGCCTCTGTTCCACATGGCCCTGGAATAATGAGGATTTAAACCAAAAATTCACACAG 840
 Qy 1363 CTGCTACTTCTACAAAGACTCAACCCAGAGGCCATGCCCTGGTCAAGAGAGCTGG 1422
 Db 841 CTGCTACTTCTACAAAGACTCAACCCAGAGGCCATGCCCTGGTCAAGAGAGCTGG 900
 Qy 1423 CTTGATTATGAGAGAGAAATCGAGACATTTAAATGAGATCAGAGACTTATTTGGCCCA 1482
 Db 901 CTTGATTATGAGAGAGAAATCGAGACATTTAAATGAGATCAGAGACTTATTTGGCCCA 960
 Qy 1483 GGCCTGGATACCAAGAACTCGCATAGTCATCTGACGGGGCTGCTGGAATTTGGGAAG 1542
 Db 961 GGCCTGGATACCAAGAACTCGCATAGTCATCTGACGGGGCTGCTGGAATTTGGGAAG 1020
 Qy 1543 TCAACACTCGCCAGGAGTGAAGGAGCCTGGGGAGAGGCCAGCTGATGGGAGCCG 1602
 Db 1021 TCAACACTCGCCAGGAGTGAAGGAGCCTGGGGAGAGGCCAGCTGATGGGAGCCG 1080
 Qy 1603 TTCCAGCATGCTTCTACTTCAGCTCAGAGAGCTGGCCAGTCCAAAGTGTGAGTCTC 1662
 Db 1081 TTCCAGCATGCTTCTACTTCAGCTCAGAGAGCTGGCCAGTCCAAAGTGTGAGTCTC 1140
 Qy 1663 GCTGAGCTCATCGGAAAGATGGGACAGCCACTCGGCTCCATTAGACAGATCCTGTCT 1722
 Db 1141 GCTGAGCTCATCGGAAAGATGGGACAGCCACTCGGCTCCATTAGACAGATCCTGTCT 1200
 Qy 1723 AGGCCAGAGCGGTCTCTTCATCTCGATGGTGTAGATGAGCCAGGATGGGTCTTGCA 1782
 Db 1201 AGGCCAGAGCGGTCTCTTCATCTCGATGGTGTAGATGAGCCAGGATGGGTCTTGCA 1260
 Qy 1783 GAGCCAGTTCAGAGCTCTGTGCACTGGAGCCAGCCACAGCCGCGGATGCACTGCTG 1842
 Db 1261 GAGCCAGTTCAGAGCTCTGTGCACTGGAGCCAGCCACAGCCGCGGATGCACTGCTG 1320
 Qy 1843 GGCAGTTCTCGGGGAAAATCTACTTCCGAGGATCCTCTGATCAGGCTCGGACC 1902
 Db 1321 GGCAGTTCTCGGGGAAAATCTACTTCCGAGGATCCTCTGATCAGGCTCGGACC 1380
 Qy 1903 ACAGCTCTGCAGAACCTCAATCTCTTTGGAGCAGGACGCTGGGTAGAGGTCCTGGGG 1962
 Db 1381 ACAGCTCTGCAGAACCTCAATCTCTTTGGAGCAGGACGCTGGGTAGAGGTCCTGGGG 1440
 Qy 1963 TTCTCTGAGTCAGAGAGGAATATTTCTACAGATATTTACAGATGAAAGGCAAGCA 2022
 Db 1441 TTCTCTGAGTCAGAGAGGAATATTTCTACAGATATTTACAGATGAAAGGCAAGCA 1500
 Qy 2023 ATTAGAGCCTTAGGTGTGTGTCATAACAAAGAGCTCTGGCCCTGTGTCTGTGCC 2082
 Db 1501 ATTAGAGCCTTAGGTGTGTGTCATAACAAAGAGCTCTGGCCCTGTGTCTGTGCC 1560
 Qy 2083 TGGGTGTCTGGCTGCTGCACTTGCTGATGACAGATGAAGCGGAGGAAATCTC 2142
 Db 1561 TGGGTGTCTGGCTGCTGCACTTGCTGATGACAGATGAAGCGGAGGAAATCTC 1620

Qy 2143 ACACCTGACTTCCAGAGCACCAACCTCTGTCTACATTACCTTGTCCAGGCTCTCCAA 2202
 Db 1621 ACACCTGACTTCCAGAGCACCAACCTCTGTCTACATTACCTTGTCCAGGCTCTCCAA 1680
 Qy 2203 GCTCAGCATTGGGACCCAGCTCAGAGCCTCTGTCTCTGGGTCTGAGGGCATCTGG 2262
 Db 1681 GCTCAGCATTGGGACCCAGCTCAGAGACCTCTGTCTCTGGGTCTGAGGGCATCTGG 1740
 Qy 2263 CAAAAAAGACCCCTTTTCAGTCCAGATGACCTCAGGAGCATGGTTAGATTGGGGCCATC 2322
 Db 1741 CAAAAAAGACCCCTTTTCAGTCCAGATGACCTCAGGAGCATGGTTAGATTGGGGCCATC 1800
 Qy 2323 ATCTCACCTCTTGAAGATGGGTATTTCTTCAAGAGCACCCCATCCCTCTGAGCTACAGC 2382
 Db 1801 ATCTCACCTCTTGAAGATGGGTATTTCTTCAAGAGCACCCCATCCCTCTGAGCTACAGC 1860
 Qy 2383 TTCAATTCACCTCTGTTTCCAGAGTCTTTTGAGCAATGTCTATGTCTTGGAGATGAG 2442
 Db 1861 TTCAATTCACCTCTGTTTCCAGAGTCTTTTGAGCAATGTCTATGTCTTGGAGATGAG 1920
 Qy 2443 AAGGGAGAGGTAAACATTTCTAATTGCATCATAGATTTGGAAAGACGCTAGAACATAT 2502
 Db 1921 AAGGGAGAGGTAAACATTTCTAATTGCATCATAGATTTGGAAAGACGCTAGAACATAT 1980
 Qy 2503 GGAATACATGGCTCTTTGGGGCATCAACACACGTTTCTTATTTGGGCTGTAAAGTGTAT 2562
 Db 1981 GGAATACATGGCTCTTTGGGGCATCAACACACGTTTCTTATTTGGGCTGTAAAGTGTAT 2040
 Qy 2563 GAGGGGAGAGAGATGGAGAAACATTTTCACTGCCGCTGTCTCAGGGAGGAAACCTG 2622
 Db 2041 GAGGGGAGAGAGATGGAGAAACATTTTCACTGCCGCTGTCTCAGGGAGGAAACCTG 2100
 Qy 2623 ATGAGTGGGTCCGCTCCCTGAGCTGTCTGAGCCACACACTCTCTGAGTCCCTCCAC 2682
 Db 2101 ATGAGTGGGTCCGCTCCCTGAGCTGTCTGAGCCACACACTCTCTGAGTCCCTCCAC 2160
 Qy 2683 TGCTTTGACGAGACTCGGAAACAAACGTTTCTTGACACAAAGTGTGGGCCATTTCCAGAA 2742
 Db 2161 TGCTTTGACGAGACTCGGAAACAAACGTTTCTTGACACAAAGTGTGGGCCATTTCCAGAA 2220
 Qy 2743 ATGGCATGTGTGTAGAAACAGACATGGAGCTTTTGTGTGCATTTCTGCATTTAAATTC 2802
 Db 2221 ATGGCATGTGTGTAGAAACAGACATGGAGCTTTTGTGTGCATTTCTGCATTTAAATTC 2280
 Qy 2803 AGCCGCCACGTGAAGAACCTTCACTGATTTAGGGCAGGACAGACAGATCAACATGGAGC 2862
 Db 2281 AGCCGCCACGTGAAGAACCTTCACTGATTTAGGGCAGGACAGACAGATCAACATGGAGC 2340
 Qy 2863 CCCACATGGTAGTCCCTGTTCAGGTGGGTCCAGTCCAGATGCCTATTGGCAGATTTCTC 2922
 Db 2341 CCCACATGGTAGTCCCTGTTCAGGTGGGTCCAGTCCAGATGCCTATTGGCAGATTTCTC 2400
 Qy 2923 TTCTCCGCTCTCAAGGTCCACAGAAACCTGAAGGAGCTGGAACCTAAGTGGAAATCTCGTG 2982
 Db 2401 TTCTCCGCTCTCAAGGTCCACAGAAACCTGAAGGAGCTGGAACCTAAGTGGAAATCTCGTG 2460
 Qy 2983 AGCCACTCTGAGGTGAAGATCTTTGTAAAGCCTTGAGACCCCTCGCTGCTGCTCTGAG 3042
 Db 2461 AGCCACTCTGAGGTGAAGATCTTTGTAAAGCCTTGAGACCCCTCGCTGCTGCTCTGAG 2520
 Qy 3043 ACCCTGCGGTGGCTGTGGCTCACAGTGAAGTCAAGGACCTTGCCTTTTGGG 3102
 Db 2521 ACCCTGCGGTGGCTGTGGCTCACAGTGAAGTCAAGGACCTTGCCTTTTGGG 2580
 Qy 3103 CTGAGAGCAACCCAGACCTTGAACGAGCTTGAACCTTCAATGTGCTCACGGATGCT 3162
 Db 2581 CTGAGAGCAACCCAGACCTTGAACGAGCTTGAACCTTCAATGTGCTCACGGATGCT 2640
 Qy 3163 GGAGCCAAACCTTTTGGCAGAGCTGAGACAGCCGAGCTGGAAGCTACAGGACTCGAG 3222
 Db 2641 GGAGCCAAACCTTTTGGCAGAGCTGAGACAGCCGAGCTGGAAGCTACAGGACTCGAG 2700


```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFP58601822"
/tissue_type="uterus"
/clone_lib="586 (synonym: hutel). Vector pSport1; host
DH10B; sites NotI + SalI/MluI"
/dev_stage="adult"
1..4123
/gene="DKFP58601822"
<1..3580
CDS
/gene="DKFP58601822"
/notes="XIA0926, partly , alternative spliced"
/codon_start=2
/product="hypothetical protein"
/protein_id="CA55945.1"
/db_xref="GI:5911939"
/translation="LLQRPHPRSQDPLVKRSPDYVENRGLHIEIRDLFGPGLDQTQ
EPRIVLQAGAGIGKSLARQVKEAWRGQLYGDHFQHFYFSCRELAQSKVSLAEI
IGKDTATPAPRIQILSRPERLLFLDGVDEPGWVLOEPSSSLCHLWSPQPADALLG
SLIKTILPEASFLITARTALONLIPSLQARWEVLGFSSESRKEYFYRYPTBERQ
AIRARFLVSNELMALCVLPWVSMALACTCLMQMKRKEKLLTSTKTTTLLCHLALQ
ALQAPLQSLDLCSLAAEGWQKTLFSPDLRKHGLDGAIIISFLKMGILQBHPH
PLSYSFILCFEFTFAAMSIVLEDEKRGKHSNCIIIEKTLIAYGIIHGLFGASTFR
LLGLLSDEGERMENIFHRLSGQRNLMOWPSLOLLQPHSLSLHLCLYEYRNTKTLF
TQVMAHFERMGCVETDMELLYCTFCIKFSRHVKLOLIEGRHRSYTSWPTWVILFRW
VPTDAYWQILSVLKVTRNLKELDLGNSLSHSYKSLCKTLRRPRCLLETFLRAGC
GUTAECKOLAFGLRANQTLIELDLSFNVLTDAGAKHLQQRQPSKLRQLQVSCG
LSDSCDDQSLASPSLKELDLQONLDDVGLLCEGLRHPACKLIRLGLDOTTLL
SDMRQELRALQOEKPOLLIIFSRKPSVMTPTGELDTGEMSNSTSLKRLQIGSRAA
SHVAQNLKLLDVSKIIFPAIEAESSEPVVVELLCPVSPASQDLHTKPLGTDDDF
WGPPTAVTEVDKENVLRVHPVAVGSVRWNTGLCFVMREAVTEVEFCVMDQFLG
EINPOHSVMVAGPLIIDKARGAENAVHLPHEVALQGGHVDTSLPQVAHKEEGMLE
KPARVELHIVLENFSFQGVLLKMHNLALFPVTSVLLYHRVHEEVTFFHLYLI
RSPGDLQFSEFYVQREGLIRLVQKDKDETLVMEALVKPGDLMPTATLIIPPARIAV
PSPDAPQLHFDVQVREQLIARVTSVEVLVDKHLGQVLSQEQYSEVRLAENTRPSQMR
KLFLSLSQSDWRCCKDGLYQALXETHPLIMELWEKSKGLLPLSS"
4057..4062
polyA_signal
/gene="DKFP58601822"
4081
polyA_site
/gene="DKFP58601822"
BASE COUNT 1002 a 1086 c 1136 g 899 t
ORIGIN

Query Match      69.9%; Score 3807.8; DB 9; Length 4123;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 3951; Conservative 0; Mismatches 2; Indels 132; Gaps 1;

QY 1365 GCTACTTCTACAAAGACCTCACCCAGAGCCNAGATCCCGTGGTCAAGAGAGCTGGCC 1424
DB 1 GCTACTTCTACAAAGACCTCACCCAGAGCCNAGATCCCGTGGTCAAGAGAGCTGGCC 60

QY 1425 TGATTATGTGGAGGAGAACTCGAGGACATTTAATTGAGATCAGAGACTTATTTGGCCGAG 1484
DB 61 TGNATTATGTGGAGGAGAACTCGAGGACATTTAATTGAGATCAGAGACTTATTTGGCCGAG 120

QY 1485 CTTGGATACCCAGAAACCTCGCATAGTCATCTGCAAGGGGGCTGCTGGAAATTTGGGAAGTC 1544
DB 121 CTTGGATACCCAGAAACCTCGCATAGTCATCTGCAAGGGGGCTGCTGGAAATTTGGGAAGTC 180

QY 1545 AACACTGGCCAGGAGTGAAGGAGCTTGGGGGAGAGCCAGCTGTATGGGGAGCCGCTT 1604
DB 181 AACACTGGCCAGGAGTGAAGGAGCTTGGGGGAGAGCCAGCTGTATGGGGAGCCGCTT 240

QY 1605 CCAGCATGTCTTCTACTTTCAGCTGTCAGAGAGCTGGCCCGCAGTCCAGAGTGGTGGTCTCGC 1664
DB 241 CCAGCATGTCTTCTACTTTCAGCTGTCAGAGAGCTGGCCCGCAGTCCAGAGTGGTGGTCTCGC 300

QY 1665 TGAGCTCATCGGAAAAGATGGGACAGCCACTCGGGCTCCCATTTAGACAGATCCTGTCTAG 1724
DB 301 TGAGCTCATCGGAAAAGATGGGACAGCCACTCGGGCTCCCATTTAGACAGATCCTGTCTAG 360

```

```

1725 GCCAGAGCGGCTGCTCTTTCATCTCGATGGTGTAGATGAGCCAGGATGGTCTTTCAGGA 1784
361 GCCAGAGCGGCTGCTCTTTCATCTCGATGGTGTAGATGAGCCAGGATGGTCTTTCAGGA 420
1785 GCCAGATTTCTAGCTCTCTCTGCACTGAGGACGAGCCAGCGGGGGATGCACTCTCTGGG 1844
421 GCCAGATTTCTAGCTCTCTCTGCACTGAGGACGAGCCAGCGGGGGATGCACTCTCTGGG 480
1845 CAGTTTTGCTGGGGGAAAACCTATATCTCCGAGGACATCTTCTCTGATCAGCGCTCGGACAC 1904
481 CAGTTTTGCTGGGGGAAAACCTATATCTCCGAGGACATCTTCTCTGATCAGCGCTCGGACAC 540
1905 AGCTCTGAGAAACCTCATTTCTTTTGGAGCAGGACAGTTGGGTAGAGGTCCTCGGGTT 1964
541 AGCTCTGAGAAACCTCATTTCTTTTGGAGCAGGACAGTTGGGTAGAGGTCCTCGGGTT 600
1965 CTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTTCAAGATGAAGGCAAGCAAT 2024
601 CTCTGAGTCCAGCAGGAAGGAATATTTCTACAGATATTTTCAAGATGAAGGCAAGCAAT 660
2025 TAGAGCCTTTAGGTTGGTCAAAACAAAAGAGCTCTGGGCCCTGTCTCTTGTGCCCTG 2084
661 TAGAGCCTTTAGGTTGGTCAAAACAAAAGAGCTCTGGGCCCTGTGTCTTGTGCCCTG 720
2085 GGTGTCTGGCTGGCTGCACTTGCTGATGAGCAGATGAAGCGGAAAGGAAAACTCAC 2144
721 GGTGTCTGGCTGGCTGCACTTGCTGATGAGCAGATGAAGCGGAAAGGAAAACTCAC 780
2145 ACTGACTTCCAAAGACCAACAAACCTCTGTCTACTTACCTTGGCCAGGCTCTCCAAGC 2204
781 ACTGACTTCCAAAGACCAACAAACCTCTGTCTACTTACCTTGGCCAGGCTCTCCAAGC 840
2205 TCAGCCATTTGGGACCCAGCTCAGAGACTCTGTCTCTCTGGCTGTGAGGGCATCTGGCA 2264
841 TCAGCCATTTGGGACCCAGCTCAGAGACTCTGTCTCTCTGGCTGTGAGGGCATCTGGCA 900
2265 AAAAAAGACCCCTTTTCAGTCCAGATGACCTCAGGAAGCATGGTTAGATGGGGCCATCAT 2324
901 AAAAAAGACCCCTTTTCAGTCCAGATGACCTCAGGAAGCATGGTTAGATGGGGCCATCAT 960
2325 CTCACCTTTTGAAGATGGGTATTTCTTCAAGAGACCCCATCTCTGAGCTACAGCTT 2384
961 CTCACCTTTTGAAGATGGGTATTTCTTCAAGAGACCCCATCTCTGAGCTACAGCTT 1020
2385 CATTCACCTCTGTTTCCAAAGAGTCTTTTTCAGCAATGCTCTATGCTTTGGAGATGAGAA 2444
1021 CATTCACCTCTGTTTCCAAAGAGTCTTTTTCAGCAATGCTCTATGCTTTGGAGATGAGAA 1080
2445 GGGGAGAGGTAAAACATTTCTAATTGTCATCATAGATTTGAAAAAGACGCTAGAGCATATGG 2504
1081 GGGGAGAGGTAAAACATTTCTAATTGTCATCATAGATTTGAAAAAGACGCTAGAGCATATGG 1140
2505 AATACATGGCTGTTTGGGGCATCAAACACGTTTCTTATTTGGGCTGTTTAAGTGATGA 2564
1141 AATACATGGCTGTTTGGGGCATCAAACACGTTTCTTATTTGGGCTGTTTAAGTGATGA 1200
2565 GGGGAGAGAGATGAGAAACATTTTCACTGCGGCTGTCTCAGGGGAGAACTCAT 2624
1201 GGGGAGAGAGATGAGAAACATTTTCACTGCGGCTGTCTCAGGGGAGAACTCAT 1260
2625 GCAGTGGGTCCCGTCCCTGCGAGCTGCTGTCAGGCCACACTCTCTGGAGTCCCTCCACTG 2684
1261 GCAGTGGGTCCCGTCCCTGCGAGCTGCTGTCAGGCCACACTCTCTGGAGTCCCTCCACTG 1320
2685 CTTGTGACAGACTCGGAAACAAAACGTTTCTGACACAAAGTATGAGCCCATTTTCGAAAGAA 2744
1321 CTTGTGACAGACTCGGAAACAAAACGTTTCTGACACAAAGTATGAGCCCATTTTCGAAAGAA 1380
2745 GGGCATGTCTGAGAAACAGACATGAGCTCTTAGTGTGCACTTTCTGCATTTAAATTCAG 2804
1381 GGGCATGTCTGAGAAACAGACATGAGCTCTTAGTGTGCACTTTCTGCATTTAAATTCAG 1440
2805 CCGCCACCTGAAAGAGCTTTCAGCTGATTGAGGGCAGGAGCAGACATCAACATGAGGCC 2864

```


Db 1441 |||||CGCCACGTAAGAAAGCTTCAGCTGATTTAGGGCAGGCGACAGATCAACATGAGGCC1500
 QY 2865 CACCATGGTAGTCCCTGTTTCAAGTGGGTCCAGTACAGATGCTATTGGCAGATTCTCTT2924
 Db 1501 CACCATGGTAGTCCCTGTTTCAAGTGGGTCCAGTACAGATGCTATTGGCAGATTCTCTT1560
 QY 2925 CTCCTGCTCAAGGTCAACAGAACTGAAGAGCTGGACCTTAAGTGGAACTCGCTGAG2984
 Db 1561 CTCCTGCTCAAGGTCAACAGAACTGAAGAGCTGGACCTTAAGTGGAACTCGCTGAG1620
 QY 2985 CCACTCTGAGTGAAGAGTCTTTGAAGACCTGAGACCCCTGAGACCCCTCGTGCCTCTGGAGAC3044
 Db 1621 CCACTCTGAGTGAAGAGTCTTTGAAGACCTGAGACCCCTGAGACCCCTCGTGCCTCTGGAGAC1680
 QY 3045 CTGCGGTGGGTGGCTGAGCTCAGCTGAGGACTGCAAGGACCTTGGCTTTGGGCT3104
 Db 1681 CTGCGGTGGGTGGCTGAGCTCAGCTGAGGACTGCAAGGACCTTGGCTTTGGGCT1740
 QY 3105 GAGAGCAACAGACCTTGACCGAGCTGGACCTGAGCTTCAATGTGCTCAACGATGCTGG3164
 Db 1741 GAGAGCAACAGACCTTGACCGAGCTGGACCTGAGCTTCAATGTGCTCAACGATGCTGG1800
 QY 3165 AGCCAAACCTTTGGCAGAGACTGAGACAGCGGAGCTGCAAGCTTACAGCGACTGAGCT3224
 Db 1801 AGCCAAACCTTTGGCAGAGACTGAGACAGCGGAGCTGCAAGCTTACAGCGACTGAGCT1860
 QY 3225 GGTGAGCTGGCTCAGCTCTGACTGCTGCTGCGAGGACCTGGCTCTGTCTTAGTGCCAG3284
 Db 1861 GGTGAGCTGGCTCAGCTCTGACTGCTGCTGCGAGGACCTGGCTCTGTCTTAGTGCCAG1920
 QY 3285 CCCCAGCTGAAGGAGCTAGACTGCGAGCAGAGCAACCTGGAGTACCTTGGCGTGGACT3344
 Db 1921 CCCCAGCTGAAGGAGCTAGACTGCGAGCAGAGCAACCTGGAGTACCTTGGCGTGGACT1980
 QY 3345 GCTCTGTAGGGGCTCAGGACTCTGCTGCGTGAACCTCATACGCTGGGCTGGACCGAGAC3404
 Db 1981 GCTCTGTAGGGGCTCAGGACTCTGCTGCGTGAACCTCATACGCTGGGCTGGACCGAGAC2040
 QY 3405 AACTCTGAGTGATGATGAGGAGGAGGAACTGAGGGCCCTGGAGGAGGAAACCTCAGCT3464
 Db 2041 AACTCTGAGTGATGATGAGGAGGAGGAACTGAGGGCCCTGGAGGAGGAAACCTCAGCT2100
 QY 3465 GCTCATCTTCAGCAGAGGAACTCAGTGTGATGACCCCTACTGAGGGGCTTGATACGGG3524
 Db 2101 GCTCATCTTCAGCAGAGGAACTCAGTGTGATGACCCCTACTGAGGGGCTTGATACGGG2160
 QY 3525 AGAGATGAGTAATAGCACATCTCTCACTAAGCGGCGAGACTCGGATCAGAGGGCGGC3584
 Db 2161 AGAGATGAGTAATAGCACATCTCTCACTAAGCGGCGAGACTCGGATCAGAGGGCGGC2220
 QY 3585 TTCCCATGTTGCTCAGGCTAATCTCAAACTCCTGGAGCTGAGCAGATCTTCCCAATTGC3644
 Db 2221 TTCCCATGTTGCTCAGGCTAATCTCAAACTCCTGGAGCTGAGCAGATCTTCCCAATTGC2280
 QY 3645 TGAGATTGAGAGGAAAGCTCCCGAGAGTGTACCGGTGGAACTCTTGTGCTGCTTC3704
 Db 2281 TGAGATTGAGAGGAAAGCTCCCGAGAGTGTACCGGTGGAACTCTTGTGCTGCTTC2340
 QY 3705 TCTGCTCTCAAGGGGACTGCGATACGAAGCTTTTGGGAGTGCAGTACTCTTGGGG3764
 Db 2341 TCTGCTCTCAAGGGGACTGCGATACGAAGCTTTTGGGAGTGCAGTACTCTTGGGG2400
 QY 3765 CCCCAGGGGCTGTGGCTACTGAGTGTGACAAAGAAAGAACTTGTACCGAGTTCA3824
 Db 2401 CCCCAGGGGCTGTGGCTACTGAGTGTGACAAAGAAAGAACTTGTACCGAGTTCA2460
 QY 3825 CTTTCCCTGTAGCTGGCTCTACCGCTGGCCCAACACGGGTCTCTGCTTGTGATGAGAGA3884
 Db 2461 CTTTCCCTGTAGCTGGCTCTACCGCTGGCCCAACACGGGTCTCTGCTTGTGATGAGAGA2520
 QY 3885 AGCGGTGACCGTTGAGATTGAATCTGTGTGGGACCAAGTTCTTGGGTGAGATCAACCC3944

Db 2521 ACCGTGACCGTTGAGATTGAATCTGTGTGGGACAGTTCTCTGGGTGAGATCAACCC2580
 QY 3945 ACAGCACAGCTGGATGGTGGCAGGGCCTCTGTGTGACATCAAGGCTGAGCCTGGAGCTGT4004
 Db 2581 ACAGCACAGCTGGATGGTGGCAGGGCCTCTGTGTGACATCAAGGCTGAGCCTGGAGCTGT2640
 QY 4005 GGAAGCTGTGACCTCCCTCACTTTGTGGCTCTCAAGGGGGCCATGTGGACACATCCCT4064
 Db 2641 GGAAGCTGTGACCTCCCTCACTTTGTGGCTCTCAAGGGGGCCATGTGGACACATCCCT2700
 QY 4065 GTTCCAAATGGCCCACTTTAAAGAGAGGGGATGCTCTGGAGAGCCAGCCAGGGTGA4124
 Db 2701 GTTCCAAATGGCCCACTTTAAAGAGAGGGGATGCTCTGGAGAGCCAGCCAGGGTGA2760
 QY 4125 GCTGATCACATAGTTCTGGAAACCCAGCTTCTCCCTTTGGGAGTCTCTCTGAAAAT4184
 Db 2761 GCTGATCACATAGTTCTGGAAACCCAGCTTCTCCCTTTGGGAGTCTCTCTGAAAAT2820
 QY 4185 GATCATTAATGCTGGCTTCAATTCCTGTCACCTCTGTGTGGTGTCTTTACACCGCT4244
 Db 2821 GATCATTAATGCTGGCTTCAATTCCTGTCACCTCTGTGTGGTGTCTTTACACCGCT2880
 QY 4245 CCATCTGAGGAGTCACTTCCACCTCTACCTGATCCCAAGTCACTCTCATTCGG--4302
 Db 2881 CCATCTGAGGAGTCACTTCCACCTCTACCTGATCCCAAGTCACTCTCATTCGGAA2940
 QY 4303 -----4302
 Db 2941 GSCCATAGATGATCTAGAAATGAAATTCAGTTTGTGGCAATCCACAGCCACCCCGCT3000
 QY 4303 -----4302
 Db 3001 GACCCCACTTATATGGGCTGTCGTACACTGTGTCTGGGTCTGGTTCAAGGATGCTGGA3060
 QY 4303 -----RAGGAATGGAGCTCTGCTATCGAAGCCCTGGAGAAGACCAAGTGTCTC4352
 Db 3061 AATACTCCCAAGGAATGGAGCTCTGTATCGAAGCCCTGGAGAAGACCAAGTGTCTC3120
 QY 4353 GGAGTTCTACGTTGGCCACTTGGGATCAGGATCAGGCTGCAAGTGAAGAAAGAAAGA4412
 Db 3121 GGAGTTCTACGTTGGCCACTTGGGATCAGGATCAGGCTGCAAGTGAAGAAAGAAAGA3180
 QY 4413 TGAGACTCTGTGTGGGAGGCTTGGTGAACCAAGAGATCTCATGCTCTGCAACTCTCT4472
 Db 3181 TGAGACTCTGTGTGGGAGGCTTGGTGAACCAAGAGATCTCATGCTCTGCAACTCTCT3240
 QY 4473 GATCCCTCAGCCCGATAGCCGTAACCTCTGAGTCCCGCCGAGTGTGCTGCACTT4532
 Db 3241 GATCCCTCAGCCCGATAGCCGTAACCTCTGAGTCCCGCCGAGTGTGCTGCACTT3300
 QY 4533 TGTGGACAGTATCGAGAGCAGCTGATAGCCCGAGTGACATCGGTGGAGGTTGTCTTGA4592
 Db 3301 TGTGGACAGTATCGAGAGCAGCTGATAGCCCGAGTGACATCGGTGGAGGTTGTCTTGA3360
 QY 4593 CAAATGCTATGACAGGCTCTGAGCAGGAGCAGTACGAGAGGCTCTGGCTGAGAACAC4652
 Db 3361 CAAATGCTATGACAGGCTCTGAGCAGGAGCAGTACGAGAGGCTCTGGCTGAGAACAC3420
 QY 4653 GAGGCCAGCAGATGCGGAAGCTGTTAGCTTGGAGTCCCTGGGACCGGAAGTGCAA4712
 Db 3421 GAGGCCAGCAGATGCGGAAGCTGTTAGCTTGGAGTCCCTGGGACCGGAAGTGCAA3480
 QY 4713 AGATGACTCTACCAAGCCCTGAAGGAGACCATCTCACTCATTTATGAACTCTGGGA4772
 Db 3481 AGATGACTCTACCAAGCCCTGAAGGAGACCATCTCTCACTCATTTATGAACTCTGGGA3540
 QY 4773 GAAGGCGAGCAAAAGGAGACTCTGCTCACTCAGCAGTGAAGTATCAACACCAAGCCCTTG4832
 Db 3541 GAAGGCGAGCAAAAGGAGACTCTGCTCACTCAGCAGTGAAGTATCAACACCAAGCCCTTG3600
 QY 4833 ACCCTTGAAGTCTGGCTTGGCTGAGCCCTTCTTGGGTCTCAGTTTCTTCTCTGAAAC4892
 Db 3601 ACCCTTGAAGTCTGGCTTGGCTGAGCCCTTCTTGGGTCTCAGTTTCTTCTCTGAAAC3660

QY 4893 AAGTTGCCATCTGGTTTGCCTTCCAGCACTAAAGTAATGAACTTTGATGATGCGCTTTGC 4952
DB 3661 AAGTTGCCATCTGGTTTGCCTTCCAGCACTAAAGTAATGAACTTTGATGATGCGCTTTGC 3720
QY 4953 TGGGCATTATGTGTCATCCAGGAGTGCACAGGGGGCCCCCAGTCCAGGTGGCCTTAACA 5012
DB 3721 TGGGCATTATGTGTCATCCAGGAGTGCACAGGGGGCCCCCAGTCCAGGTGGCCTTAACA 3780
QY 5013 GCATCTCAGGGAATGTCATCTGGAGCTGGCAAGACCCCTGCAGACCTCATAGAGCTCA 5072
DB 3781 GCATCTCAGGGAATGTCATCTGGAGCTGGCAAGACCCCTGCAGACCTCATAGAGCTCA 3840
QY 5073 TCTGTTGGCCACAGAGCAAGCCTAGAGCCCTCGGATCCCATCCAGGCGCAAGAGGA 5132
DB 3841 TCTGTTGGCCACAGAGCAAGCCTAGAGCCCTCGGATCCCATCCAGGCGCAAGAGGA 3900
QY 5133 ATAGGAGGACATGGAACCATTTGCTCTGGCTGTGTCAAGGGTGGAGCCCCCAAAATGG 5192
DB 3901 ATAGGAGGACATGGAACCATTTGCTCTGGCTGTGTCAAGGGTGGAGCCCCCAAAATGG 3960
QY 5193 GGTTCAGCTGGAGGCCACAGTGGATCTTGGCTTTGTGACAGAGATCTCAAGAGCAA 5252
DB 3961 GGTTCAGCTGGAGGCCACAGTGGATCTTGGCTTTGTGACAGAGATCTCAAGAGCAA 4020
QY 5253 GCCACAGAGTAAGTGGAGGAAGTTTATTAGAAAAATAAGGAGTATCAGAGCTCTTT 5312
DB 4021 GCCACAGAGTAAGTGGAGGAAGTTTATTAGAAAAATAAGGAGTATCAGAGCTCTTT 4080
QY 5313 TAGAA 5317
DB 4081 TAAAA 4085

RESULT 15
LOCUS AX089769
DEFINITION Sequence 9 from Patent WO0116170.
ACCESSION AX089769
VERSION AX089769.1 GI:13443941
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.

REFERENCE 1
AUTHORS Reed, J.C.
TITLE Card proteins involved in cell death regulation
JOURNAL Patent: WO 0116170-A 9 08-MAR-2001;
The Burnham Institute (US)

FEATURES
source
1..4556
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
1..4365
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC34903.1"
/db_xref="GI:13443942"

CDS
1..4556
/translation="MAGGAWRLACVLEPLKKEELKELFOLLANKAHSRSSGETPAQ
PKTSGMEVASLVIAQYGEQRANDLALHTWEOGLRSLCAQAGGAGHSPSPSPSE
PHLGSFQSTAVLMPWHELHPAGTCQSERVLRQLPDTSGRWRREISALLYQAL
PSSPDHSPQSNAPSTAVLGSWSPQSLAPROEAPGTPWLDSTGIIYYTE
IREREREKGRPWAAVVGTPPQAHTSLOPHHPWEPVRESLCTWPNKNEFNQ
KFTQLLLQRPDRSDPLVRSWPDYVEENRHLIEIRDLFGPLDTPQPRIVILQ
AAGIKSTLAROVKEAMGRGOLYGRFQHVYFSCRELAQSKVSLAEIIGKDGATP
APIRQLSPERLLFILDOVDEPGWVLPSPSELCLHWSQPQPADALLGLSLKTLIP
EASFLLTARTALQNLIPLEQARWVEVLGFSSESRKEYFYRIFIDERQAIRFLVK
SNKELWALCVLPWNSLACTCLMQMKRKEKLTLSKTTTLLCLHYLAQQAQPLHL
QLRDLCSLAAGIWKQTLFSDPLRKHLGDGAIISTFLMKILQIHPILPSYFIHL
CFQEFPAAMSVLEDEKGRGHNSNCIIIDLEKTLAYGIHLFGASTFRLGLLSDGE
BREMENI FHLRSLQGRNLWQVPSLQLLLQLPHSLHCLYETRNKTFITQVMAFEE

BASE COUNT 1100 a 1223 c 1256 g 977 t
ORIGIN
Query Match 62.5%; Score 3400.4; DB 6; Length 4556;
Best Local Similarity 93.7%; Pred. No. 0;
Matches 3555; Conservative 0; Mismatches .236; Indels 3; Gaps 1;
QY 523 ATGGCTGGCGGAGCTGGGGCGGCTGGCTGTACTTGGAGTTCCTGAAGAGGAGGAG 582
DB 1 ATGGCTGGCGGAGCTGGGGCGGCTGGCTGTACTTGGAGTTCCTGAAGAGGAGGAG 60
QY 583 CTGAAGAGGATTCAGCTTCTGCTCGCCCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG 642
DB 61 CTGAAGAGGATTCAGCTTCTGCTCGCCCAATAAAGCGCACTCCAGGAGCTCTTCGGGTGAG 120
QY 643 ACACCCGCTCAGCCAGAGAAAGACGAGTGGCATGAGAGTGGCTCTGATCTGGTGGCTCAG 702
DB 121 ACACCCGCTCAGCCAGAGAAAGACGAGTGGCATGAGAGTGGCTCTGATCTGGTGGCTCAG 180
QY 703 TATGGGGAGCAGCGGCGCTGGGACCTAGCCCTCCATACCTGGGACAGATGGGGCTGAGG 762
DB 181 TATGGGGAGCAGCGGCGCTGGGACCTAGCCCTCCATACCTGGGACAGATGGGGCTGAGG 240
QY 763 TCACCTGTGGCCCAAGCCAGAGGGGAGGCCACTCTCCCTCATTTCCCTTACAGCCCA 822
DB 241 TCACCTGTGGCCCAAGCCAGAGGGGAGGCCACTCTCCCTCATTTCCCTTACAGCCCA 300
QY 823 AGTGAACCCCACTGGGGTCTCCAGCCAAACCACTCCACCGCATGTGTAATGCCCTGG 882
DB 301 AGTGAACCCCACTGGGGTCTCCAGCCAAACCACTCCACCGCATGTGTAATGCCCTGG 360
QY 883 ATCCATGAATTCGCGCGGGTGCACCCAGGCTCAGAGAGAGAGGTTTTCAGACAGCTG 942
DB 361 ATCCATGAATTCGCGCGGGTGCACCCAGGCTCAGAGAGAGAGGTTTTCAGACAGCTG 420
QY 943 CCTGACACATCTGGACGCGCTGGAGAGAAATCTCTGCTCACTCTCTACCAAGCTCTT 1002
DB 421 CCTGACACATCTGGACGCGCTGGAGAGAAATCTCTGCTCACTCTCTACCAAGCTCTT 480
QY 1003 CCAAGCTCCCAGACCATGAGTCTCAAGCCAGGAGTCAACCAACGCCGCCCATCTCCACA 1062
DB 481 CCAAGCTCCCAGACCATGAGTCTCAAGCCAGGAGTCAACCAACGCCGCCCATCTCCACA 540
QY 1063 GCAGTGTGGGAGCTGGGATCCCCACCTCAGCCCGCTAGCACCCAGAGAGGAGGAG 1122
DB 541 GCAGTGTGGGAGCTGGGATCCCCACCTCAGCCCGCTAGCACCCAGAGAGGAGGAG 600
QY 1123 GCTCTGGGACCCCAATGGCTCTCGATGAAACGTCAGAAATTTACTACAGAAATTCAGA 1182
DB 601 GCTCTGGGACCCCAATGGCTCTCGATGAAACGTCAGAAATTTACTACAGAAATTCAGA 660
QY 1183 GAAAGAGAGAGAGAAATTCAGAGAAAGAGCGAGGCCCTCCATGGCGACGGTGTAGAACG 1242
DB 661 GAAAGAGAGAGAGAAATTCAGAGAAAGAGCGAGGCCCTCCATGGCGACGGTGTAGAACG 720
QY 1243 CCCCCACAGGCGCACACAGGCTCAGGCCCCACACCCATCCAGGAGCTCTCTGTGAGA 1302
DB 721 CCCCCACAGGCGCACACAGGCTCAGGCCCCACACCCATCCAGGAGCTCTCTGTGAGA 780
QY 1303 GAGAGCTCTGTTCCACATGGCCCTGGGAAATAGGATTTTAAACAAAAATTCACACAG 1362

QY 3523 GGAGAGATGAGTAATAGCACATCCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGCG 3582
 Db |||||
 QY 3001 GGAGAGATGAGTAATAGCACATCCTCACTCAAGCGGCAGAGACTCGGATCAGAGAGGCG 3060
 Db |||||
 QY 3583 GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCCTGGACGTGAGCAAGATCTTCCCAATT 3642
 Db |||||
 QY 3061 GCTTCCCATGTTGCTCAGGCTAATCTCAAACTCCTGGACGTGAGCAAGATCTTCCCAATT 3120
 Db |||||
 QY 3643 GCTGAGATTGCAGAGAAAGCTCCCGAGAGGTAGTACCGGTGGAACTCTTGTGCGTGCCT 3702
 Db |||||
 QY 3121 GCTGAGATTGCAGAGAAAGCTCCCGAGAGGTAGTACCGGTGGAACTCTTGTGCGTGCCT 3180
 Db |||||
 QY 3703 TCTCTGCTCTCTCAAGGGACCTGCAATACGAAGCCTTTGGGGACTGACGATGACTTCTGG 3762
 Db |||||
 QY 3181 TCTCTGCTCTCTCAAGGGACCTGCAATACGAAGCCTTTGGGGACTGACGATGACTTCTGG 3240
 Db |||||
 QY 3763 GGGCCCAACGGGCGCTGTGCTACTGAGGTAGTTGACAAAGAAAGAACTTGTACCGAGTT 3822
 Db |||||
 QY 3241 GGGCCTGAAGGAAATGTGATGTTGATTTGATTAAGAGCACAAACAGATACAGCGTT 3300
 Db |||||
 QY 3823 CACTTCCCTGTAGTGGCTCTACCGCTGGCCCAACACGGGTCTCTGCTTTGTGATGAGA 3882
 Db |||||
 QY 3301 TGGTTCCTCCACTGCTGGCTGGTATCTGTGTGAGTCAAGAGGCTGGCTTCTGTGTAAGG 3360
 Db |||||
 QY 3883 GAAGCGGTGACCGTTGAGATTGAAATCTGTGTGGGACAGTTCCTGG- --GTGAGATC 3939
 Db |||||
 QY 3361 GATGAGTCAAGTGCAGATTGGGTTGGTTCTTGGAGTCAGACCTGGCCCTGGACCTG 3420
 Db |||||
 QY 3940 AACCACACACAGCTGGATGTTGGGAGGGGCTCTGCTGGACATCAAGGCTGAGCCTGGA 3999
 Db |||||
 QY 3421 CAGCACCATGAAACAGTGGCTGTGGGCGGCCCTTGTGATGTCATGCAAGAGCCAGAG 3480
 Db |||||
 QY 4000 GCTGTGGAAGCTGTGACCTCCTCACTTTGTGGCTCTCAAGGGGGCCATGTGGACACA 4059
 Db |||||
 QY 3481 GAGGCTGTCCCGGAAATCCACCTCCCGCACTTCATCTCCCTCCAAAGGTGAGGTGGACGTC 3540
 Db |||||
 QY 4060 TCCCTGTTCCTCAATGCCCCACTTTAAAGAGGAGGGGATGCTCTGGAGAAAGCCAGCCAGG 4119
 Db |||||
 QY 3541 TCCCTGTTTCTCGTTGCCATTTTAAAGATGAAGGATGGTCTGGAGCATCCAGCCCGG 3600
 Db |||||
 QY 4120 GTGGAGCTGCATACATAGTTCTTGGAAAAACCCAGCTTCTCCCGCTTGGGAGTCTCCTG 4179
 Db |||||
 QY 3601 GTGGAGCCTTTCTATGCTGTCTGGAAAGCCCGAGCTTCTCTGATGGGCATCCTGCTG 3660
 Db |||||
 QY 4180 AAAATGATCCATTAATGCCCTGGCTTCATTCCGTCACCTCTGTGTGTGCTTTTACACAC 4239
 Db |||||
 QY 3661 CGGATCGCCAGTGGGACTCGCCTCTCCATCCCGCATCACTTCCAAACACATTGATCTATTAT 3720
 Db |||||
 QY 4240 CGGCTCCATCCTGAGGAAGTCACTTTCCACCTCTACCTGATCCCAAGTCACTGCTCCATT 4299
 Db |||||
 QY 3721 CACCCCGACCCGAGATATTAAAGTTCCACTTGTAACCTTGTCCCGAGGACGCTTGCTA 3780
 Db |||||
 QY 4300 CGGAAGGAATGGA 4313
 Db |||||
 QY 3781 ACAAGGCGATAGA 3794
 Db |||||

Search completed: January 29, 2004, 09:40:33
 Job time : 19613.5 secs